NHH

INSTITUTT FOR FORETAKSØKONOMI

DEPARTMENT OF BUSINESS AND MANAGEMENT SCIENCE

FOR 2 2013
ISSN: 1500-4066
May 2013

Discussion paper

Recursive utility and disappearing puzzles for continuous-time models

BY
Knut K. Aase

Norges Handelshøyskole

NORWEGIAN SCHOOL OF ECONOMICS

Recursive utility and disappearing puzzles for continuous-time models

Knut K. Aase *

May 10, 2013

Abstract

Motivated by the problems of the conventional model in rationalizing market data, we derive the equilibrium interest rate and risk premiums using recursive utility in a continuous time model. Two ordinally equivalent versions are considered. The state price is not Markov in any of the versions, so instead of using dynamic programming we use the stochastic maximum principle. The resulting equilibriums are consistent with low values of the parameters of the utility functions when calibrated to market data. One version is consistent with preference for early resolution of uncertainty, the other for late for the US-data. We therefore consider heterogeneity with recursive utilities. Our resulting model rationalize data well, and can explain both the Equity Premium Puzzle and the Risk-Free Rate Puzzle with good margins.

KEYWORDS: The equity premium puzzle, the risk-free rate puzzle, recursive utility, utility gradients, the stochastic maximum principle, heterogeneity, limited market participation , optimal asset allocation
JEL-Code: G10, G12, D9, D51, D53, D90, E21.

---

*The Norwegian School of Economics and Business Administration, 5045 Bergen Norway and Centre of Mathematics for Applications (CMA), University of Oslo, Norway. Telephone: (+47) 55959249. E-mail: Knut.Aase@NHH.NO. The paper was presented at the international conference "The Social Discount Rate" held in Bergen in May, 2012, and organized by Kåre Petter Hagen in cooperation with the Ministry of Finance, Norway. Special thanks to Thore Johnsen, Steinar Ekern, Gunnar Eskeland, Darrell Duffie, Bjørn Eraker, John Campbell, Bernt Øksendal and Rajnish Mehra for valuable comments. Any remaining errors are mine.

1 Introduction

Rational expectations, a cornerstone of modern economics and finance, has been under attack for quite some time. Are asset prices too volatile relative to the information arriving in the market? Is the mean risk premium on equities over the riskless rate too large? Is the real interest rate too low? Is the market's risk aversion too high?

Mehra and Prescott (1985) raised some of these questions in their well-known paper, using a variation of Lucas's (1978) pure exchange economy with a Kydland and Prescott (1982) "calibration" exercise. They chose the parameters of the endowment process to match the sample mean, variance and the annual growth rate of per capita consumption in the years 1889 - 1978. The puzzle is that they were unable to find a plausible parameter pair of the utility discount rate and the relative risk aversion to match the sample mean of the annual real rate of interest and of the equity premium over the 90-year period.

The puzzle has been verified by many others, e.g., Hansen and Singleton (1983), Ferson (1983), Grossman, Melino, and Shiller (1987). Many theories have been suggested during the years to explain the puzzle, but to date there does not seem to be any consensus that the puzzles have been fully resolved by any single of the proposed explanations [1].

We utilize a continuous time setting, to take full advantage of the analytic power of infinite dimensional analysis. We use the framework established by Duffie and Epstein (1992a-b) and Duffie and Skiadas (1994) which elaborates the foundational work by Kreps and Porteus (1978) of recursive utility in dynamic models. This latter model is extended to a continuous time setting, where future utility is a conditional expected time integral of a felicity index minus a measure of Arrow-Pratt absolute risk aversion multiplied by the variance rate of utility. When there is no uncertainty, the felicity index does not depend upon risk aversion.

While Duffie and Epstein (1992a) use dynamic programming to find risk premiums using a richer economic environment that we have, we employ the stochastic maximum principle. An alternative is to use directional derivatives

[1] Constantinides (1990) introduced habit persistence in the preferences of the agents. Also Campbell and Cochrane (1999) used habit formation. Rietz (1988) introduced financial catastrophes, Barro (2005) developed this further, Weil (1992) introduced non-diversifiable background risk, and Heaton and Lucas (1996) introduce transaction costs. There is a rather long list of other approaches aimed to solve the puzzles, among them are borrowing constraints (Constantinides et al. (2001)), taxes (Mc Grattan and Prescott (2003)), loss aversion (Benartzi and Thaler (1995)), survivorship bias (Brown, Goetzmann and Ross (1995)), and heavy tails and parameter uncertainty (Weitzmann (2007)).

and utility gradients. Neither of these principles require the Markov property. Our model for the financial market is standard, but imply no assumptions about normal returns and lognormal prices. Also Epstein and Zin (1989-91) base their results on dynamic programming in a diserete-time framework.

When calibrated to market data we find that the representative agent in an ordinally equivalent version of our model prefers *early* to late resolution of uncertainty, for plausible values of the parameters. The other, nonordinal version of the model is consistent with preference for *late* resolution of uncertainty for US 90-year dataset we employ. In another economy with a higher volatility of the market portfolio this picture is turned around. To our knowledge, this unordinal version has not been explored earlier in the continuous-time setting.

We therefore suggest a heterogeneous model where the representative agent is composed by the two ordinally equivalent versions. The resulting model is found to explain well both the Equity Premium Puzzle as well as the Risk-Free Rate Puzzle, and seems as a good description of many economies. This model is also used to discuss the non-participating issue, as well as optimal asset allocation.

Recursive preferences deviate from the separable time additive case in several important ways, and it is not at all clear that dynamic programming, an important tool for the conventional model, will give the correct answers. To cite David Kreps (1988, p 175)

" ..when the uncertainty of the gambles being optimized over resolves at dates in the future (after important decisions left out must be taken), then use of the standard models is *very* suspect and often quite wrong."

This seems to be the root of these puzzles. We know that the standard model does not fit well real data. In the conventional model, however, dynamic programming works. In the new model, with recursive utility, researchers have continued to use dynamic programming.

For recursive utility uncertainty is "dated" by the time of its resolution, and the individual regards uncertainties resolving at different times as being different. For such a complex representation the dynamic programming approach may be too restrictive, at least for the application to the market data that we have in mind.

As is well known recursive utility leads to separation of risk aversion from the elasticity of intertemporal substitution in consumption, within a time-consistent model framework. We adhere to the principle of Ockham's razor, namely to change only one single feature of the standard model. We demonstrate that going from the standard time separable and additive expected utility representation to recursive utility is sufficient for the resulting model to rationalize data well.

We have a similar analysis in discrete-time (Aase (2013), that also give rise to two ordinally equivalent versions of recursive utility, which explain the data as in the present paper. The discrete-time analysis gives a clear indication why our results are different from those based on dynamic programming, including why Weil (1989) found that recursive utility leads to even larger values for the risk-free interest rate than the conventional model (the so-called Risk-Free Rate Puzzle).

The respective risk premiums in the discrete-time version are analogous to our present results, only the interest rates differ. In the discrete-time models there is dependence on the consumption-wealth ratio in the expressions for the interest rate, which does not occur in our framework. Since the continuous-time model does not rely on approximations, it is likely to give the best representation of the data.

In addition to giving new insights about these interconnected puzzles, our model is likely to yield many other results that are difficult, or impossible, to obtain using the conventional approach. One example is that it can explain the empirical regularities for Government bills, and also for optimal asset allocation.

For recursive utility in continuous time the volatility of the future utility plays an important role. We show that this quantity is a linear combination of the volatility of the market portfolio and the volatility of aggregate consumption, by market clearing in the financial market, and by using basic properties of the recursive utility function. It is at this point the two ordinally equivalent versions differ.

There is by now a long standing literature that has been utilizing recursive preferences. We mention Avramov and Hore (2007), Avramov et al. (2010), Eraker and Shaliastovich (2009), Hansen, Heaton, Lee, Roussanov (2007), Hansen and Scheinkman (2009), Wacther (2012), Bansal and Yaron (2004), Campbell (1996), Bansal and Yaron (2004), Kocherlakota (1990 b), and Ai (2012) to name some important contributions. Related work is also in Browning et al. (1999), and on consumption see Attanasio (1999). A few exceptions to late resolution exist in this literature. Bansal and Yaron (2004) study a richer economic environment than we employ, as is typical of most of the newer literature using recursive utility.

The paper is organized as follows: In Section 2 we explain the problems with the conventional model, and give a preview of the results, focusing on the ordinally equivalent version. In Section 3 we present a brief introduction to recursive utility along the lines of Duffie and Epstein (1992a-b) and Duffie and Skiadas (1994). In Section 4 we derive the first order conditions of optimal consumption for both versions of recursive utility, where we employ the stochastic maximum principle. In Section 5 we derive equilibrium risk premiums and the interest rate for the ordinally equivalent model. In Section 6 we we present the model for the financial market, and connect the volatility of the future utility process, a quantity in the representation of preferences in our approach, to a linear function of the volatility of the market portfolio and the consumption growth rate. In Section 7 we discuss the results for the ordinally equivalent model. In Section 8 we do the analysis for the nonordinal model. In Section 9 w we solve the aggregation problem over agents, with the two different versions of recursive utility. Here we address the limited market participation issue. Section 10 treats the optimal asset allocation problem, Section 11 points out extensions, and Section 12 concludes.

2 The problems with the standard model

2.1 The additive and separable Eu-model

The conventional asset pricing model in financial economics, the consumption-based capital asset pricing model (CCAPM) of Lucas (1978) and Breeden (1979), assumes a representative agent with a utility function of consumption that is the expectation of a sum, or a time integral, of future discounted utility functions. The model has been criticized for several reasons. First, it does not perform well empirically. Second, the standard specification of utility can not separate the risk aversion from the elasticity of intertemporal substitution, while it would clearly be advantageous to disentangle these two conceptually different aspects of preference. Third, while this representation seems to function well in deterministic settings, and for *timeless* situations, it is not well founded for *temporal* problems (e.g., derived preferences may not satisfy the substitution axiom (Mossin (1969)).

In the conventional model the utility $U(c)$ of a consumption stream c_t is given by

$$U(c) = E\Big\{ \int_0^T u(c_t, t)\, dt \Big\} \tag{1}$$

where the felicity index u has the separable form

$$u(c,t) = \frac{1}{1-\gamma} c^{1-\gamma}\, e^{-\beta\, t}. \tag{2}$$

The parameter γ is the representative agent's relative risk aversion and β is the utility discount rate, or the impatience rate, and T is the time horizon. These parameters are assumed to satisfy $\gamma > 0$, $\beta \geq 0$, and $T \leq \infty$.

In this model the risk premium $(\mu_R - r)$ of any risky security can be shown to have the simple form

$$\mu_R(t) - r_t = \gamma\, \sigma_{Rc}(t) \tag{3}$$

where r_t is the equilibrium real interest rate at time t, and the term $\sigma_{Rc}(t) = \sum_{i=1}^{d} \sigma_{R,i}(t)\sigma_{c,i}(t)$ is, by the Ito-isometry, the covariance rate between returns of the risky asset and the growth rate of aggregate consumption at time t, a measurable and adaptive process satisfying standard conditions. The dimension of the Brownian motion is $d > 1$. This is the continuous-time version of Breeden's consumption-based CAPM. Similarly, the expression for the risk-free real interest rate is

$$r_t = \beta + \gamma\, \mu_c(t) - \frac{1}{2}\gamma\,(\gamma + 1)\, \sigma_c'(t)\sigma_c(t). \tag{4}$$

The process $\mu_c(t)$ is the annual growth rate of aggregate consumption and $(\sigma_c'(t)\sigma_c(t))$ is the annual variance rate of consumption growths, both at time t, again dictated by the Ito-isometry. Both these quantities are measurable and adaptive stochastic processes, satisfying standard conditions. The return processes as well as the consumption growth rate process in this paper are also assumed to be ergodic processes, implying that statistical estimation makes sense.

Notice that in the model is the instantaneous correlation coefficient between returns and the consumption growth rate given by

$$\kappa_{Rc}(t) = \frac{\sigma_{Rc}(t)}{||\sigma_R(t)|| \cdot ||\sigma_c(t)||} = \frac{\sum_{i=1}^{d} \sigma_{R,i}(t)\sigma_{c,i}(t)}{\sqrt{\sum_{i=1}^{d} \sigma_{R,i}(t)^2}\sqrt{\sum_{i=1}^{d} \sigma_{c,i}(t)^2}},$$

and similarly for other correlations given in this model. Here $-1 \leq \kappa_{Rc}(t) \leq 1$ for all t. Note that with this convention we can equally well write $\sigma_R(t)\sigma_c(t)$ for $\sigma_{Rc}(t)$, and the former does *not* imply that the instantaneous correlation coefficient between returns and the consumption growth rate is equal to one.

In Table 1 we reproduce from Mehra and Prescott (1985) the key summary statistics of the real annual return data related to the S&P-500, denoted by M, as well as for the annualized consumption data, denoted c, and the government bills, denoted b. [2].

Here we have, for example, estimated the covariance between aggregate consumption and the stock index directly from the data set to be .00223. This gives the estimate .3752 for the correlation coefficient.[3]

[2]There are of course newer data by now, but these retain the same basic features. If we can explain the data in Table 1, we can explain any of the newer sets as well.

[3]The full data set was provided by Professor Rajnish Mehra.

	Expectat.	Standard dev.	Covariances
Consumption growth	1.83%	3.57%	$\text{cov}(M, c) = .002226$
Return S&P-500	6.98%	16.54%	$\text{cov}(M, b) = .001401$
Government bills	0.80%	5.67%	$\text{cov}(c, b) = -.000158$
Equity premium	6.18%	16.67%	

Table 1: Key US-data for the time period 1889 -1978. Discrete-time compounding.

	Expectation	Standard dev.	Covariances
Consumption growth	1.81%	3.55%	$\hat{\sigma}_{Mc} = .002268$
Return S&P-500	6.78%	15.84%	$\hat{\sigma}_{Mb} = .001477$
Government bills	0.80%	5.74%	$\hat{\sigma}_{cb} = -.000149$
Equity premium	5.98%	15.95%	

Table 2: Key US-data for the time period 1889 -1978. Continuous-time compounding.

Since our development is in continuous time, we have carried out standard adjustments for continuous-time compounding, from discrete-time compounding. The results of these operations are presented in Table 2. This gives, e.g., the estimate $\hat{\kappa}_{Mc} = .4033$ for the instantaneous correlation coefficient $\kappa(t)$. The overall changes are in principle small, and do not influence our comparisons to any significant degree, but are still important.

Interpreting the risky asset as the value weighted market portfolio M corresponding to the S&P-500 index, we have two equation in two unknowns to provide estimates for the preference parameters by the "method of moments". Indeed, what we really do here is to use the assumption about ergodicity of the various μ_t and σ_t processes. This enables us to replace "state averages" by "time averages", the latter being given in Table 2. The result is

$$\gamma = 26.37 \quad \beta = -.015,$$

i.e., a relative risk aversion of about 26 and an impatience rate of minus 1.5%.

If we insist on a nonnegative impatience rate, as we probably should (but see Kocherlakota (1990)), this means that the real interest rate explained by the model is larger than 3.3% (when $\beta = .01$, say) for the period considered, but it is estimated, as is seen from Table 2, to be less than one per cent.

We denote the elasticity of intertemporal substitution in consumption by ψ, and refer to it as the EIS-parameter. In the standard model $\psi = 1/\gamma$, so if the risk aversion is as large as indicated in the above, it means that $\psi = .037$, which is considered to be too low for the average individual.

2.2 Including Government bills

There is also another problem with the standard model. From Table 2 we see that there is a negative correlation between Government bills and the consumption growth rate. Similarly there is a positive correlation between the return on S&P-500 and Government bills.

If we interpret Government bills as risk free, the former correlation should be zero for the CCAPM-model to be consistent. Since this correlation is not zero, then $\gamma = 0$ if Government bills are to be risk free, which is inconsistent with the model. Since the Government bills used by Mehra and Prescott (1985) have duration one month, and the data are yearly, Government bills are not the same as Sovereign bonds with duration of one year. One month bills in a yearly context will then contain price risk 11 months each year, and hence the real risk free rate should, perhaps, be set strictly lower that .80%. Assuming the risk premium of Government bills is about fourth of a per cent (.0040) and using the expression for the equilibrium risk premium (3) also for the bills,

$$\mu_b(t) - r_t = \gamma\, \sigma_{cb}(t), \tag{5}$$

we solve this equation together with the equation for the real risk free interest rate (now estimated to .0040), using the estimate of $\sigma_{cb}(t)$ in Table 2. This gives the calibrated values $\gamma = -23.84$ and $\beta = .93$, also inconsistent with the above calibrated values (and very implausible).

Another version could be to insert the estimate for γ obtained above, 26.37, in this equation together with the estimate for $\sigma_{cb}(t)$, which is $-.000149$ from Table 2. This gives a risk premium for Government bills of $-.0039$, which is also implausible.

On the other hand, whatever positive value for the risk premium we choose, the resulting value of γ is negative. With bills included, the standard model does not seem to have enough 'degrees of freedom' to match the data, since in this situation the model contains three basic relationships and only two 'free parameters'. Thus the standard model can not have the correct state price deflator.

To better understand the problems with contemporary asset pricing, we propose to consider recursive utility along the lines of Duffie and Epstein (1992a-b), where risk aversion and consumption substitution can be separated. This will have clear implications for risk premiums and the equilibrium interest rate, as we shall demonstrate.

2.3 Preview of our results

Let ρ the *time preference* of the individual. Our approach allows $\rho = 1/\psi$ to be different from risk aversion γ. Based on the analysis to be presented in Section 6, the two relationships corresponding to (3) and (4) are, for $\rho \neq 1$ and with the same notation as above, given as follows:

$$\mu_R(t) - r_t = \frac{\rho(1-\gamma)}{1-\rho}\, \sigma'_R(t)\sigma_c(t) + \frac{\gamma-\rho}{1-\rho}\, \sigma'_R(t)\sigma_M(t) \qquad (6)$$

and

$$r_t = \beta + \rho\mu_c(t) - \frac{1}{2}\,\frac{\rho(1-\rho\gamma)}{1-\rho}\, \sigma'_c(t)\sigma_c(t) + \frac{1}{2}\,\frac{\rho-\gamma}{1-\rho}\, \sigma'_M(t)\sigma_M(t) \qquad (7)$$

respectively.

The risk premiums in (6) are endogenously derived in (41) and the same is true for the expression for the equilibrium interest rate in (42). Here $\sigma_M(t)$ signifies the volatility of the return on the market portfolio of the risky securities, $\sigma'_R(t)\sigma_M(t) = \sigma_{RM}(t)$ is the instantaneous covariance rate of the returns on any risky asset labeled R with the return of the market portfolio. In the model these quantities are assumed to be measurable, adaptive, ergodic stochastic processes satisfying standard conditions.

As can be readily seen from these two expressions a larger risk premium can be explained by this model when $\gamma > \rho$ and $\rho < 1$, while at the same time the real interest rate may then be smaller than for the standard model. This could explain the two most serious problems with the latter model.

This form of the model implies that it can be estimated by linear regression.

The risk premium of any risky asset in (6) is seen to be a linear combination of the market-based CAPM of Mossin (1966) and the consumption-based CAPM of Breeden (1979). If $\gamma = \rho$ risk premiums reduce to those of the latter.

In particular all the μ and σ- processes are assumed to be *ergodic* so that it makes sense to estimate them, as explained in the previous section. Note that there is no contradiction in having these quantities assessed by numerical values, although they are stochastic processes in the theory [4].

Indeed, we here make use of our assumption about ergodicity of the drift and diffusion terms: These are defined as conditional expectations, i.e.,

[4] If these quantities are assumed to be constants also in the theoretical model, this would lead to the usual contradiction, where the volatility of the market portfolio would have to be the same as the volatility in the consumption growth. That this is not so is evident from the summary statistics in Table 2.

stochastic processes, and by ergodicity we may replace state averages by time averages[5]

When calibrating we fix the time impatience rate β to some reasonable number, say .027, and solve the two non-linear equations (6) and (7) using the data summarized in Table 2, when when the risky security is the market index ($R = M$). The results of this are

$$\gamma = 1.50 \quad \text{and} \quad \rho = 0.66 \quad \text{corresponding to} \quad EIS = 1.51.$$

as the estimates for the remaining two parameters in the recursive utility function.

The numerical values of these parameters should not be taken literary, but rather as an indication that this model rationalizes the data well. Nevertheless, a short discussion where values of this magnitude are taken as given is of interest.

The relatively low value of the time preference ρ a representative agent who does not require too much compensation for consumption substitution in a deterministic world. This value of γ must be considered as low for the relative risk aversion (in particular compared to 26 for the standard model). An impatience rate of 2.7% is considered acceptable (in particular compared to -1.5%).

In other words, with these values of the preference parameters of the recursive-utility-representative-agent, the model can explain an equilibrium interest rate and equity premium estimated to, respectively

$$\hat{r} = .0080 \quad \text{and} \quad (\hat{\mu}_M - \hat{r}) = .0598 \quad \text{(continuous time adjusted values)}$$

for the consumption/market data used by Mehra and Prescott (1985), presented in Table 2. This is a solution of both the Equity Premium Puzzle of Mehra and Prescott (1985) as well as the Risk-Free Rate Puzzle of Weil (1989).

That the risk premium can be large in our model is illustrated by the market based CAPM term, when $\gamma > \rho$ and $\rho < 1$, explaining the "missing link" of the risk premium in the CCAPM-specification. The richer model allows a reconciliation of the data in Table 2.

One challenge with the conventional model is that it provides too high interest rates if β is constrained to be non-negative. It is lower in our model for several intuitive reasons: The second term in (7) containing μ_c does not contribute nearly as much any more any more since a reasonably low value of ρ replaces a large value of γ. The "precautionary savings" term works in the

[5]The spirit of this part is also contained in the research of Haavelmo (1944).

right direction provided $\rho < 1$ and $\rho\gamma < 1$. If $\rho < 1$ and $\rho\gamma > 1$, but the ratio $\rho(1-\gamma\rho)/(1-\rho)$ is not to large, a small interest rate can result from the last term provided $\gamma > \rho$ and $\rho < 1$, both of which we find plausible. Together this may explain an equilibrium real interest rate of less than one per cent for the data summarized in Table 2, and with very reasonable parameters. At the same time this may explain the large, observed equity risk premium for a representative agent who favors early to late resolution of uncertainty ($\gamma > \rho$).

Notice that the term corresponding to precautionary savings is negative provided $1 > \gamma\rho$ and $\rho < 1$, which is then one requirement for the recursive utility consumer to be 'prudent'. The other is $1 < \gamma\rho$ and $\rho > 1$.

If ρ is constrained to be zero, the model reduces to

$$\mu_R(t) - r_t = \gamma\, \sigma_{RM}(t), \qquad r_t = \beta - \frac{\gamma}{2}\sigma'_M(t)\sigma_M(t).$$

The risk premium is that of the ordinary CAPM-type, while the interest rate is new. This version of the model corresponds to "neutrality" of consumption transfers in some sense, to be explained later. Solving these two non-linear equations consistent with the data of Table 2, we obtain

$$\gamma = 2.38 \quad \text{and} \quad \beta = .038.$$

In the conventional model this simply gives risk neutrality, i.e., $\gamma = \rho = 0$, so this model gives a risk premium of zero, and a short rate of $r = \beta$.

Figure 1 illustrates the the feasible region in (ρ, γ)-space. For the conventional model it is the 45°-line shown ($\rho = \gamma$). For the recursive utility model it is all of the first quadrant, including the axes. The points above the 45°-line represent late resolution of uncertainty, the points below correspond to early resolution. As can be seen, both the point ($\rho = .66, \gamma = 1.50, \beta = .027$) reported above, denoted Calibr 1, and the point corresponding to the market based CAPM, called CAPM++ in the figure, ($\rho = 0, \gamma = 2.38, \beta = .038$), are located in the *early resolution* part. The point denoted Calibr 3 is ($\gamma = 1/2$, $\rho = 1.29$ and $\beta = .01$) and corresponds to late resolution of uncertainty, for the utility function $h(x) = \sqrt{x}$. The point denoted Calibr 2 is associated with the nonordinal model and will be explained later.

Estimates of the EIS-parameter seem difficult to obtain for several reasons, and the results will naturally depend on circumstances. In e.g., Dagsvik et al. (2006) an estimate of this parameter is suggested to be in the range from 1 to 1.5.

The larger region for the (ρ, γ)-combinations permitted by our model is not a frivolous generalization of the conventional model. Numerous generalizations have been presented without achieving fully acceptable resolutions. That the richer structure of the recursive model is a modest extension

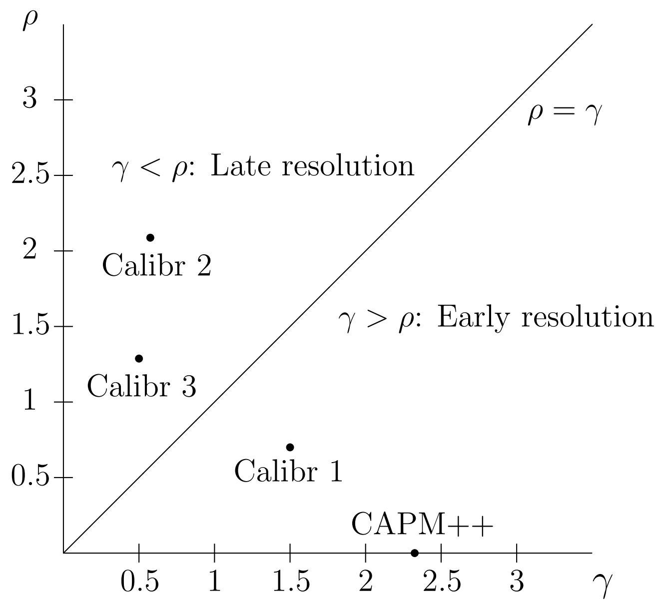

Figure 1: Calibration points in the (γ, ρ)-space

is demonstrated by the interpretations and plausible results yielded in our simple expressions. It is based on fundamental assumptions and axioms of rational behavior.

The risk premium of any risky asset is seen to depend on other risky assets through the volatility of the market portfolio, and the return rate on government bonds depends both on how aggregate consumption covariates with the stock market as well as the size of the variance of the market portfolio.

Initially one would think that these features should be reflected also in the corresponding formulas in the conventional model, but at the outset it is hard to say if these aspects are internalized or not.

2.4 Fitting to the data including Government bills

In the above discussion we have interpreted Government bills as risk free. As mentioned, this may not be entirely correct. Exactly what risk premium bills command we can here only stipulate. With the same assumption as in the previous section, for a risk premium of .0040 for the bills we have a third equation, namely

$$\mu_b(t) - r_t = \frac{\rho(1-\gamma)}{1-\rho}\,\sigma_{b,c}(t) + \frac{\gamma-\rho}{1-\rho}\,\sigma_{b,M} \tag{8}$$

to solve together with the equations (6) and (7). With the covariance estimates provided in Table 2, we have three equations in three unknowns, giving the following solution

$$\beta = .027, \quad \gamma = 1.76 \quad \text{and} \quad \rho = .53 \quad (EIS = 1.97).$$

The results of this calibration are somewhat sensitive to changes in the risk premium for the Government bills.

Referring to the situation called CAPM++ discussed above, where we constrained ρ to be zero, equation (7) and (8) give $\gamma = 2.71$ and $\beta = .038$ for this situation. While this is not exactly the same values ($\gamma = 2.38$, $\beta = .038$) as the ones obtained from the equations (6) and (7) with $\rho = 0$, it is not far off in this regard (compare to the standard model of subsection 2.1).

3 Recursive Stochastic Differentiable Utility

In this section we recall the essentials of recursive, stochastic, differentiable utility along the lines of Duffie and Epstein (1992a-b) and Duffie and Skiadas (1994).

Despite the fact that the analysis naturally becomes more technically involved once we depart from the additive and separable framework of the expected utility representation, we obtain surprisingly simple and transparent results when we use the Kreps-Porteus specification for the felicity index. The issue of when uncertainty is resolved is an important one in this theory, as Figure 1 illustrates.

We are given a probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, t \in [0,T], P)$ satisfying the 'usual' conditions, and a standard model for the stock market with Brownian motion driven uncertainty, N risky securities and one riskless asset (Section 6 provides more details). Consumption processes are chosen form the space L of square integrable progressively measurable processes with values in R_+.

The stochastic differential utility $U : L \to R$ is defined as follows by two primitive functions: $f : L \times R \to R$ and $A : R \to R$.

The function $f(c_t, V_t)$ is a felicity index at time t, and A is a measure of absolute risk aversion of the Arrow-Pratt type for the agent. In addition to current consumption c_t, the felicity index also depends on future utility.

The utility process V for a given consumption process c, satisfying $V_T = 0$, is given by the representation

$$V_t = E_t\Big\{ \int_t^T \big(f(c_s, V_s) - \frac{1}{2} A(V_s)\, \tilde{\sigma}_V(s)'\tilde{\sigma}_V(s)\big) ds\Big\}, \quad t \in [0, T] \qquad (9)$$

where E_t denotes conditional expectation given $\mathcal{F}_t$ and $\tilde{\sigma}_V(t)$ is an R^d-valued square-integrable progressively measurable volatility process. Here d is the dimension of the Brownian motion B_t. We think of V_t as the remaining utility for c at time t, conditional on current information $\mathcal{F}_t$, and $A(V_t)$ is penalizing for risk.

If, for each consumption process c_t, there is a well-defined utility process V, the stochastic differential utility U is defined by $U(c) = V_0$, the initial utility. The pair (f, A) generating V is called an aggregator.

Since $V_T = 0$ and $\int \tilde{\sigma}_V(t) dB_t$ is a martingale, (9) has the stochastic differential equation representation

$$dV_t = \Big(- f(c_t, V_t) + \frac{1}{2} A(V_t)\, \tilde{\sigma}_V(t)'\tilde{\sigma}_V(t)\Big)\, dt + \tilde{\sigma}_V(t)\, dB_t \qquad (10)$$

If terminal utility different from zero is of interest, like for life insurance, then V_T may be different from zero. We think of A as associated with a function $h : R \to R$ such that $A(v) = -\frac{h''(v)}{h'(v)}$, where h is two times continuously differentiable. U is monotonic and risk averse if $A(\cdot) \geq 0$ and f is jointly concave and increasing in consumption [6].

The representation (9) is motivated from the discrete time model. The common starting point for recursive utility is that future utility at time t is given by $V_t = W(c_t, m(V_{t+1}))$ for some function W, where m is a certainty equivalent at time t. If h is a von Neumann-Morgenstern index, then $m(V) = h^{-1}(E[h(V)])$ and the preferences fall into the Kreps and Porteus (1978) family. The certainty equivalent m is then assumed to satisfy some smoothness properties. With Brownian information V_t is an Ito process, and based on this Duffie and Epstein (1992b) demonstrate how (9) is justified.

Stochastic differential utility partly disentangles intertemporal substitution from risk aversion: In the case of deterministic consumption, $\tilde{\sigma}_V(t) = 0$ a.s. for all t. Hence risk aversion A is then irrelevant, since it multiplies a zero variance. Thus certainty preferences, including the willingness to substitute consumption across time, are determined by f alone. Only risk attitudes are affected by changes in A for f fixed. In particular, if

$$\tilde{A}(\cdot) \geq A(\cdot)$$

[6]In the general case $A(\cdot)$ is associated with a local gradient representation (LGR) $M(v, x)$ of a certainty equivalent m. When $m = h^{-1}(E[h(V)])$ for h a von Neumann-Morgenstern index, then $A(V_t) = -M_{1,1}(V_t, V_t)$ where $M_{1,1}(v, x) = \partial^2 M(v, x)/\partial v^2$.

where U and $\tilde{U}$ are utility functions corresponding to (f, A) and $(f, \tilde{A})$ respectively, then $\tilde{U}$ is more risk averse than U in the sense that any consumption process c rejected by U in favor of some deterministic process $\bar{c}$ would also be rejected by $\tilde{U}$. Thus

$$U(c) \leq U(\bar{c}) \Rightarrow \tilde{U}(c) \leq \tilde{U}(\bar{c}). \tag{11}$$

Here it is important that $f(c_t, V_t)$ at the outset does not depend on risk aversion, only on time substitution.

The preference ordering represented by recursive utility is *time consistent* in the sense of Johnsen and Donaldson (1985).

In this paper we consider two ordinally equivalent specifications: The first has the Kreps-Porteus utility representation, which corresponds to the aggregator with a CES specification

$$f_1(c, v) = \frac{\beta}{1-\rho} \frac{c^{1-\rho} - v^{1-\rho}}{v^{-\rho}} \quad \text{and} \quad A_1(v) = \frac{\gamma}{v} \tag{12}$$

corresponding to two utility functions $u(c) = \frac{c^{1-\rho}}{1-\rho}$ and $h(v) = \frac{v^{1-\gamma}}{1-\gamma}$, say. If, for example, $A_1(v) = 0$ for all v, this means that the recursive utility agent is risk neutral.

Here $\rho \geq 0, \rho \neq 1, \beta \geq 0, \gamma \geq 0, \gamma \neq 1$ (when $\rho = 1$ or $\gamma = 1$ it is the logarithms that apply). The elasticity of intertemporal substitution in consumption $\psi = 1/\rho$. The parameter ρ is the time preference parameter referred to in Section 2.2. Here $u(\cdot)$ and $h(\cdot)$ are different functions, resulting in the desired disentangling of γ from ρ.

An ordinally equivalent specification is the following: When the aggregator (f_1, A_1) is given corresponding to the utility function U_1, there exists a strictly increasing and smooth function $\varphi(\cdot)$ such that the ordinally equivalent $U_2 = \varphi \circ U_1$ has the aggregator (f_2, A_2) where

$$f_2(c, v) = ((1-\gamma)v)^{-\frac{\gamma}{1-\gamma}} f_1(c, ((1-\gamma)v)^{\frac{1}{1-\gamma}}), \qquad A_2 = 0,$$

and

$$U_2 = \frac{1}{1-\gamma} U_1^{1-\gamma}. \tag{13}$$

This is the second specification we work with, and it has has the CES-form

$$f_2(c, v) = \frac{\beta}{1-\rho} \frac{c^{1-\rho} - ((1-\gamma)v)^{\frac{1-\rho}{1-\gamma}}}{((1-\gamma)v)^{\frac{1-\rho}{1-\gamma}-1}}, \qquad \tilde{A}_2(v) = 0. \tag{14}$$

Is should be emphasized that the reduction to a normalized aggregator $(f_2, 0)$ does not mean that intertemporal utility is risk neutral, or that we have lost the ability to separate risk aversion from substitution (see Duffie and Epstein(1992a)). The corresponding utility U_2 retains the essential features, namely that of (partly) disentangling intertemporal elasticity of substitution from risk aversion.

Here it is instructive to recall the that the standard additive and separable utility has aggregator

$$f(c, v) = u(c) - \beta v, \quad A = 0. \tag{15}$$

in the present framework (an ordinally equivalent one). As can be seen, even if $A = 0$, the agent of the standard model is not risk neutral.

4 The First Order Conditions

In the following we look at the solution of the above problems. For any of the versions $i = 1, 2$ formulated in the previous section, the representative agent's problem is to solve

$$\sup_{\tilde{c} \in L} U(\tilde{c})$$

subject to

$$E\Big\{\int_0^T \tilde{c}_t \pi_t dt\Big\} \leq E\Big\{\int_0^T c_t \pi_t dt\Big\}.$$

Here $V_t = V_t^{\tilde{c}}$ is the solution of the backward stochastic differential equation (BSDE)

$$\begin{cases} dV_t = -f(t, \tilde{c}_t, V_t, \tilde{\sigma}_V(t))\, dt + \tilde{\sigma}_V(t)\, dB_t \\ V_T = 0. \end{cases} \tag{16}$$

Notice that (16) covers both the versions (12) and (14) that we intend to analyze. For $\alpha > 0$ define the Lagrangian

$$\mathcal{L}(\tilde{c}; \lambda) = U(\tilde{c}) - \alpha\, E\Big(\int_0^T \pi_t(\tilde{c}_t - c_t) dt\Big)$$

In order to find the first order condition for the representative consumer's problem, we use Kuhn-Tucker and either directional derivatives in function space, or the stochastic maximum principle. Neither of these principles require the Markov property. The problem is well posed since U is increasing and concave and the constraint is convex. In maximizing the Lagrangian of the problem, we can calculate the directional derivative $\bigtriangledown U(c; h)$, which equals $(\bigtriangledown U(c))(h)$ where $\bigtriangledown U(c)$ is the gradient of U at c. Since U is continuously differentiable, this gradient is a linear and continuous functional, and thus, by the Riesz representation theorem, it is given by an inner product.

Because of the generality of the problem, let us instead utilize the stochastic maximum principle (see Pontryagin (1972), Bismut (1978), Kushner (1972), Bensoussan (1983), or Peng (1990)) : We then have a forward backward stochastic differential equation (FBSDE) system consisting of the simple FSDE $dX(t) = 0; X(0) = 0$ and the BSDE (16). The objective functional is

$$J(\tilde{c}) = V_0^{\tilde{c}} - \alpha\, E\Big(\int_0^T \pi_t(\tilde{c}_t - c_t) dt \Big) \tag{17}$$

where α is the Lagrange multiplier. The Hamiltonian for this problem is

$$H(t, \tilde{c}, v, \tilde{\sigma}_v, y) = y_t\, f(t, \tilde{c}_t, v_t, \tilde{\sigma}_v(t)) - \alpha\, \pi_t(\tilde{c}_t - c_t) \tag{18}$$

and the adjoint equation is

$$\begin{cases} dY_t = Y(t)\big(\frac{\partial f}{\partial v}(t, \tilde{c}_t, V_t, \tilde{\sigma}_V(t))\, dt + \frac{\partial f}{\partial z}(t, \tilde{c}_t, V_t, \tilde{\sigma}_V(t))\, dB_t\big) \\ Y_0 = 1. \end{cases} \tag{19}$$

where we have used the notation $Z(t) = \tilde{\sigma}_V(t)$, and z as the generic variable. If c^* is optimal we therefore have

$$Y_t = \exp\Big(\int_0^t \{ \frac{\partial f}{\partial v}(s, c_s^*, V_s, \tilde{\sigma}_V(s)) - \frac{1}{2}\big(\frac{\partial f}{\partial z}(s, c_s^*, V_s, \tilde{\sigma}_V(s))\big)^2 \}\, ds + \int_0^t \frac{\partial f}{\partial z}(s, c_s^*, V_s, \tilde{\sigma}_V(s))\, dB(s) \Big) \quad a.s. \tag{20}$$

Maximizing the Hamiltonian with respect to $\tilde{c}$ gives the first order equation

$$y\, \frac{\partial f}{\partial \tilde{c}}(t, c^*, v, \tilde{\sigma}_v) - \alpha\, \pi = 0$$

or

$$\alpha\, \pi_t = Y(t) \frac{\partial f}{\partial \tilde{c}}(t, c_t^*, V(t), \tilde{\sigma}_V(t)) \quad \text{a.s. for all } t \in [0, T]. \tag{21}$$

Notice that the state price deflator π_t at time t depends, through the term Y_t, on the entire, optimal paths $(c_s, V_s, \tilde{\sigma}_V(t))$ for $0 \le s \le t$, which means that in general $\pi(t)$ will not be a Markov process. This is the strength of the stochastic maximum principle; it does not require the Markov property.

When $\gamma = \rho$ then $Y_t = e^{-\beta t}$ for the aggregator (15) of the conventional model, so the state price deflator is a Markov process, and dynamic programming (DP) is a possible technique to apply. If $\gamma \neq \rho$ on the other hand, π_t is *not* a Markov process, and the requirements for DP are not satisfied. Instead we have uses directional derivatives in function space in the time-discrete version, and the stochastic maximum principle in the continuous-time model of this paper.

For the representative agent equilibrium the optimal consumption process is the given aggregate consumption c in society, and for this consumption process the remaining utility V_t at time t is optimal.

We now have the first order conditions for both the versions of recursive utility outlined in Section 3. We start with the ordinally equivalent version denoted 2 with aggregator given by (14).

5 The analysis for the ordinally equivalent model

For this model the first order conditions are given by

$$\alpha\, \pi_t = Y_t\, \frac{\partial f}{\partial c}(c_t, V_t) \qquad \text{a.s. for all } t \in [0, T] \tag{22}$$

where $f(t, c, v, \tilde{\sigma}_v) = f_2(c, v)$ is given in (14), and where the adjoint variable $Y(t)$ is

$$Y_t = \exp\Big(\int_0^t \frac{\partial f}{\partial v}(c_s, V_s)\, ds \Big) \quad a.s. \tag{23}$$

As can be noted, for this version the adjoint process is of bounded variation[7].

Aggregate consumption is exogenous, with dynamics on of the form

$$\frac{dc_t}{c_t} = \mu_c(t)\, dt + \sigma_c(t)\, dB_t, \tag{24}$$

where $\mu_c(t)$ and $\sigma_c(t)$ are measurable, $\mathcal{F}_t$ adapted stochastic processes, satisfying appropriate integrability properties. This is also assumed for processes representing returns. In addition we assume these processes to be ergodic, so that we may replace time averages by state averages.

Similarly the process V_t is assumed to follow the dynamics

$$\frac{dV_t}{(1-\gamma)V_t} = \mu_V(t)\, dt + \sigma_V(t)\, dB_t \tag{25}$$

where

$$\tilde{\sigma}_V(t) = (1-\gamma)V_t \sigma_V(t), \quad \text{and} \quad \mu_V(t) = -\frac{\beta}{1-\rho}\Big(\frac{c_t^{1-\rho} - ((1-\gamma)V_t)^{\frac{1-\rho}{1-\gamma}}}{((1-\gamma)V_t)^{\frac{1-\rho}{1-\rho}}} \Big).$$

[7]Originally the author derived this FOC using utility gradients based on a result of Duffie and Skiadas (1994).

When $\gamma > 1$ utility V is negative so the product $(1-\gamma)V > 0$ a.s., which gives us a positive volatility of V provided $\sigma_V(t) > 0$ a.e. From the FOC (43) we then get the dynamics of the state price deflator:

$$d\pi_t = f_c(c_t, V_t)\, dY_t + Y_t\, df_c(c_t, V_t). \tag{26}$$

Using Ito's lemma this becomes

$$\begin{aligned} d\pi_t =\ & Y_t\, f_c(c_t, V_t)\, f_v(c_t, V_t)\, dt + Y_t\, \frac{\partial f_c}{\partial c}(c_t, V_t)\, dc_t + Y_t\, \frac{\partial f_c}{\partial v}(c_t, V_t)\, dV_t \\ & + Y_t\Big(\frac{1}{2}\,\frac{\partial^2 f_c}{\partial c^2}(c_t, V_t)\,(dc_t)^2 + \frac{\partial^2 f_c}{\partial c\,\partial v}(c_t, V_t)\,(dc_t)(dV_t) + \frac{1}{2}\,\frac{\partial^2 f_c}{\partial v^2}(c_t, V_t)\,(dV_t)^2\Big). \end{aligned} \tag{27}$$

Here

$$f_c(c, v) := \frac{\partial f(c, v)}{\partial c} = \frac{\beta\, c^{-\rho}}{\big((1-\gamma)v\big)^{\frac{1-\rho}{1-\gamma}-1}},$$

$$f_v(c, v) := \frac{\partial f(c, v)}{\partial v} = \frac{\beta}{1-\rho}\ \Big(c^{1-\rho}\big((1-\gamma)v\big)^{-\frac{1-\rho}{1-\gamma}}(\rho-\gamma) + (\gamma-1)\Big),$$

$$\frac{\partial f_c(c, v)}{\partial c} = -\frac{\beta\,\rho\, c^{-\rho-1}}{\big((1-\gamma)v\big)^{\frac{\gamma-\rho}{1-\gamma}}}, \qquad \frac{\partial f_c(c, v)}{\partial v} = \beta(\rho-\gamma)\, c^{-\rho}\,\big((1-\gamma)v\big)^{-\frac{1-\rho}{1-\gamma}},$$

$$\frac{\partial^2 f_c}{\partial c^2}(c, v) = \frac{\beta\,\rho\,(1+\rho)\, c^{-\rho-2}}{\big((1-\gamma)v\big)^{\frac{1-\rho}{1-\gamma}-1}}, \qquad \frac{\partial^2 f_c}{\partial c\,\partial v}(c, v) = \frac{\rho\,\beta\,(\gamma-\rho)\, c^{-\rho-1}}{\big((1-\gamma)\big)^{\frac{1-\rho}{1-\gamma}}},$$

and

$$\frac{\partial^2 f_c}{\partial v^2}(c, v) = \frac{\beta\,(\gamma-\rho)\,(1-\rho)\, c^{-\rho}}{\big((1-\gamma)v\big)^{\frac{1-\rho}{1-\gamma}+1}}.$$

Denoting the dynamics of the state price deflator by

$$d\pi_t = \mu_\pi(t)\, dt + \sigma_\pi(t)\, dB_t, \tag{28}$$

from (27) and the above expressions we now have that the drift and the diffusion terms of π_t are given by

$$\begin{aligned} \mu_\pi(t) = Y_t\Big(& \frac{\beta^2}{1-\rho}(\rho-\gamma)\, c_t^{2(1-\rho)-1}((1-\gamma)V_t)^{-\frac{2(1-\rho)}{1-\gamma}+1} \\ & - \frac{(1-\gamma)\beta^2}{1-\rho}\, c_t^{-\rho}((1-\gamma)V_t)^{-\frac{1-\rho}{1-\gamma}+1} - \beta\,\rho\, c_t^{-\rho}((1-\gamma)V_t)^{-\frac{1-\rho}{1-\gamma}+1}\,\mu_c(t) \\ & -\beta\, c_t^{-\rho}\,(\rho-\gamma)\,((1-\gamma)V_t)^{-\frac{1-\rho}{1-\gamma}}\, f(c_t, V_t) + \frac{1}{2}\,\beta\,\rho\,(1+\rho)\, c_t^{-\rho}\,((1-\gamma)V_t)^{-\frac{1-\rho}{1-\gamma}+1}\,\sigma_c'(t)\sigma_c(t) \\ & - \beta\,\rho c_t^{-\rho}\,(\rho-\gamma)\,((1-\gamma)V_t)^{-\frac{1-\rho}{1-\gamma}+1}\sigma_{cV}(t) \\ & - \frac{1}{2}\,\beta\,(\rho-\gamma)\,(1-\rho)\, c_t^{-\rho}\,((1-\gamma)V_t)^{-\frac{1-\rho}{1-\gamma}+1}\,\sigma_V'(t)\sigma_V(t)\Big), \end{aligned} \tag{29}$$

and

$$\sigma_\pi(t) = Y_t\,\beta\,c_t^{-\rho}\Big((-\rho)\,\sigma_c(t)\,((1-\gamma)V_t)^{-\frac{1-\rho}{1-\gamma}+1}+ \\ (\rho-\gamma)\,((1-\gamma)V_t)^{-\frac{1-\rho}{1-\gamma}}((1-\gamma)V_t)\sigma_V(t)\Big) \quad (30)$$

respectively.

5.1 The risk premium for the ordinally equivalent version

The risk premium is in general given by

$$\mu_R(t) - r_t = -\frac{1}{\pi_t}\,\sigma_{R\pi}(t), \quad (31)$$

where $\sigma_{R\pi}(t)$ is the instantaneous covariance of the increments of R and π. Interpreting π_t as the price of the consumption good at time t, by the first order condition it is a decreasing function of consumption c since $f_{cc} < 0$.

Combining the FOC with the result in (30), the formula for the risk premium in terms of the primitives of the model is accordingly given by

$$\mu_R(t) - r_t = \rho\,\sigma_{Rc}(t) + (\gamma-\rho)\,\sigma_{RV}(t). \quad (32)$$

This is a very basic result of our analysis, and turns out to be the same also for the nonordinal version based on (12), as we demonstrate later. This result can be seen to point to a potential solution to the Equity Premium Puzzle: Provided the covariance rate between the index (R = M) and future utility V is positive (and reasonable), the latter term on the right-hand side of (32) is the candidate to fill the gap between the observed (large) risk premium and the one explained by the CCAPM term.

Observe that when $\rho = 0$ the last term explains all of the risk premium, in which case the utility function u of consumption in the CES-specification of the felicity index f is of the form $u(c) = c$. This corresponds to neutrality with respect to consumption transfers.

When $\rho \neq \gamma$ the latter term in (32) may be positive or negative. It turns out that the most reasonable situation for the data summarized in Table 2 is when $\gamma > \rho$ for the present version, corresponding to early resolution of uncertainty. This results in a higher equilibrium risk premium produced by the recursive utility model than for the conventional model, as we shall demonstrate.

We return to the equilibrium determination of the volatility term $\sigma_V(t)$. Before we do that, we give an expression for the equilibrium interest rate r_t, also in terms of $\sigma_V(t)$.

5.2 The equilibrium interest rate for the ordinally equivalent version

The equilibrium interest rate r_t is given by the general formula

$$r_t = -\frac{\mu_\pi(t)}{\pi_t}. \tag{33}$$

The real interest rate at time t can be thought of as the expected exponential rate of decline of the representative agent's marginal utility, which is π_t in equilibrium.

In order to find an expression for r_t in terms of the primitives of the model, we use the formula for $f(c_t, V_t)$ from (14) in the expression for $\mu_\pi(t)$ in (29). We then obtain the following

$$\begin{aligned} r_t = \beta + \rho\mu_c(t) - \frac{1}{2}\rho\big(\rho+1\big)\sigma_c'(t)\sigma_c(t) + \\ \rho(\rho-\gamma)\sigma_{cV}(t) + \frac{1}{2}(\rho-\gamma)(1-\rho)\sigma_V'(t)\sigma_V(t). \end{aligned} \tag{34}$$

This is the second basic result of our analysis, and turns out to be the same also for the nonordinal version based on (12) as well. This result can be seen to indicate a potential solution to the Risk-Free Rate Puzzle: Provided the covariance rate between the aggregate consumption growth rate and future utility V is positive (and reasonable), the fourth term on the right-hand side of (34) may lower the model interest rate relative to the standard model provided $\gamma > \rho$. The latter term may also work in the same direction provided $\rho < 1$. Again, these relationships between the parameters seem reasonable.

For the standard utility ($\rho = \gamma$) this reduces the interest rate to the familiar expression in (4). We observe that it is the time substitution interpretation that is the meaningful one in this new setting for terms two and three on the right hand side. First and foremost it turns out to be the term related to the growth rate of consumption that now contributes to a lower value of r, because the parameter ρ turns out to calibrate to a reasonably low value. The "precautionary savings" term also works in the right direction since it is negative, but is likely to be relatively small in magnitude. The two last terms are likely to be negative provided $\rho < \gamma$ and $\rho \leq 1$.

In order to link the volatility term $\sigma_V(t)$ to an observable (or estimable) quantity in the market, we now specify a model for the financial market.

6 A simple model for the financial market

Our model for the financial economy is standard, and merely a time-continuous version of the model employed by Mehra and Prescott (1985). Thus we do not change anything related to the standard Lucas model except the preference structure.

Let $\nu(t) \in R^N$ denote the vector of expected rates of return of the N given risky securities in excess of the riskless instantaneous return r_t, and let $\sigma(t)$ denote the matrix of diffusion coefficients of the risky asset prices, normalized by the asset prices, so that $\sigma(t)\sigma(t)'$ is the instantaneous covariance matrix for asset returns. Both $\nu(t)$ and $\sigma(t)$ are progressively measurable, ergodic processes.

The representative consumer's problem is, for each initial level w of wealth to solve

$$\sup_{(c,\varphi)} U(c) \tag{35}$$

subject to the intertemporal budget constraint

$$dW_t = \big(W_t(\varphi_t' \cdot \nu(t) + r_t) - c_t\big)dt + W_t\varphi_t' \cdot \sigma(t)dB_t, \tag{36}$$

Here $\varphi_t' = (\varphi_t^{(1)}, \varphi_t^{(2)}, \cdots, \varphi_t^{(N)})$ are the fractions of total wealth W_t held in the risky securities.

Market clearing requires that $(\varphi_t)'\sigma(t) = (\delta_t^M)'\sigma(t) = \sigma_M(t)$ in equilibrium, where $\sigma_M(t)$ is the volatility of the return on the market portfolio, and δ_t^M are the fractions of the different securities, $j = 1, \cdots, N$ held in the value-weighted market portfolio. That is, the representative agent can only hold the market portfolio in equilibrium, by definition.

6.1 The volatility of future utility for the ordinally equivalent model

We recall the following: Recursive utility of the Kreps and Porteus type that we use is *homothetic*, meaning that for any consumption processes c and c' and any scalar $\lambda > 0$

$$U(\lambda c') \geq U(\lambda c) \Leftrightarrow U(c') \geq U(c).$$

For our version of recursive utility this translates to (see Duffie and Epstein (1992b): An aggregator (f_2, A_2) that generates a recursive utility function U_2 is homothetic if there exists an odinally equivalent aggregator (f_1, A_1) satisfying (i) f_1 is homogeneous of degree 1, and (ii) and the variance multiplier A_1 is linearly homogeneous of degree -1, i.e., there is some k such that $A_1(v) = k/v$ for all $v > 0$. The nonordinal utility function U_1 generated by the (f_1, A_1) of Section 3 is homogeneous of degree one, that is $U_1(\lambda c) = \lambda U_1(c)$ for all c and $\lambda > 0$.

Recall the connection between the two ordinally equivalent recursive utility representations that we deal with. It follows from (13) that

$$\begin{aligned} \bigtriangledown U_2(c^*; c^*) &= U_1(c)^{-\gamma} \bigtriangledown U_1(c^*; c^*) \\ &= U_1(c^*)^{-\gamma} U_1(c^*) = U_1(c^*)^{1-\gamma} = (1-\gamma) U_2(c^*). \end{aligned} \quad (37)$$

The second equality follows, since by the definition of directional derivatives

$$\begin{aligned} \bigtriangledown U_1(c^*; c^*) &= \lim_{\alpha \downarrow 0} \frac{U_1(c^* + \alpha c^*) - U_1(c^*)}{\alpha} = \lim_{\alpha \downarrow 0} \frac{U_1(c^*(1+\alpha)) - U_1(c^*)}{\alpha} \\ &= \lim_{\alpha \downarrow 0} \frac{(1+\alpha) U_1(c^*) - U_1(c^*)}{\alpha} = \lim_{\alpha \downarrow 0} \frac{\alpha U_1(c^*)}{\alpha} = U_1(c^*) \end{aligned}$$

where the third equality uses that U is homogeneous of degree one. It follows from the first order condition that

$$\bigtriangledown \tilde{U}(c^*; c^*) = E\Big(\int_0^T \pi_t c_t^* \, dt \Big) = W_0 \pi_0$$

where W_0 is the wealth of the representative agent at time zero. Let $\tilde{V}_t(c_t^*)$ and $V_t(c_t^*)$ denote future utility at the optimal consumption for our representation and the nonordinal version of recursive utility, respectively. The same, basic relationship holds here for the associated directional derivatives, i.e.,

$$\bigtriangledown \tilde{V}_t(c^*; c^*) = E_t\Big(\int_t^T \pi_s c_s^* \, ds \Big) = W_t \pi_t = V_t^{1-\gamma} = (1-\gamma) \tilde{V}_t.$$

This shows that

$$\tilde{V}_t := \tilde{V}_t(c_t^*) = \frac{\pi_t W_t}{1-\gamma} \quad (38)$$

so that for recursive utility the optimal future utility is a function of both wealth and the state price deflator π_t. Since π_t is not a Markov process, dynamic programming does not seem appropriate.

This relationship says that the optimal future utility is proportional to the nominal value of current wealth.

Note that the wealth at any time t is given by

$$W_t = \frac{1}{\pi_t} E_t\Big(\int_t^T \pi_s c_s^* \, ds \Big),$$

and thus depends on past consumption and utility from time zero to the present time t.

From (36) it follows that $\sigma_W(t) = W_t\sigma_M(t)$, and since the Ito process $\tilde{V}_t$ is a function of the agent's wealth and the state price deflator, it is a consequence of Ito's lemma that its diffusion term is

$$\tilde{\sigma}_V(t) = \tilde{V}_w(t)W_t\sigma_M(t) + \tilde{V}_\pi(t)\sigma_\pi(t) \tag{39}$$

where $\tilde{\sigma}_V(t)$ is the diffusion term of $\tilde{V}$. By (25) we have the following relationship

$$\tilde{\sigma}_V(t) = \sigma_V(t)(1-\gamma)\tilde{V}_t,$$

and form (38) it follows that $\tilde{V}_w(t) = \pi_t/(1-\gamma)$ and $\tilde{V}_\pi(t) = W_t/(1-\gamma)$. Altogether this gives

$$\tilde{\sigma}_V(t) = \frac{\pi_t W_t}{1-\gamma}\,\sigma_M(t) + \frac{\pi_t W_t}{1-\gamma}\,\frac{\sigma_\pi(t)}{\pi_t}.$$

We now use the representations for for π_t and $\sigma_\pi(t)$ in (30)-(32) to solve this problem. This gives

$$\sigma_V(t)(1-\gamma)\tilde{V}_t = \frac{\pi_t W_t}{1-\gamma}\Big(\sigma_M(t) - \rho\sigma_c(t) + (\rho-\gamma)\sigma_V(t)\Big).$$

By (38) we are left with

$$\sigma_V(t)(1-\gamma) = \sigma_M(t) - \rho\sigma_c(t) + (\rho-\gamma)\sigma_V(t),$$

from which we can determine the unknown $\sigma_V(t)$. This normalized volatility is accordingly given by

$$\sigma_V(t) = \frac{1}{1-\rho}\,\Big(\sigma_M(t) - \rho\sigma_c(t)\Big). \tag{40}$$

In the expressions for equilibrium risk premiums and the real interest rate this was the only unknown quantity. Inserting (40) into (32) and (33) we finally obtain the expressions

$$\mu_R(t) - r_t = \frac{\rho(1-\gamma)}{1-\rho}\,\sigma_R(t)\sigma_c(t) + \frac{\gamma-\rho}{1-\rho}\,\sigma_R(t)\sigma_M(t), \tag{41}$$

and

$$r_t = \beta + \rho\mu_c(t) - \frac{1}{2}\,\frac{\rho(1-\rho\gamma)}{1-\rho}\,\sigma_c(t)\sigma_c(t) + \frac{1}{2}\,\frac{\rho-\gamma}{1-\rho}\,\sigma_M(t)\sigma_M(t). \tag{42}$$

Using dynamic programming with our model for a financial market does not lead to the basic representations is Sections 5.1 and 5.2. In order to determine the volatility of future utility $\sigma_V(t)$, the dynamic programming approach produces a different result from the one we have found. For the simple case of a constant investment opportunity set we have found an exact solution to the generalized Bellman equation for recursive utility. It says that the optimal remaining utility as a function of wealth alone, the indirect utility function, is of the form $J(W_t,t) = h(t)w^\theta/\theta$ for some deterministic function $h(\cdot)$ and constant θ. Here $\theta = (1-\gamma)$. This implies that $\tilde{\sigma}_V(t) = \sigma_J(t) = J_w W_t\, \sigma_M(t)$, and using the above it follows that

$$\sigma_V(t) = \frac{\tilde{\sigma}_V(t)}{\theta V_t} = \frac{J_w W_t\, \sigma_M}{\theta J(W(t))} = \frac{h(t) W_t^{\theta-1} W_t\, \sigma_M(t)}{h(t) W_t^\theta} = \sigma_M(t)$$

which is not consistent with (40). Solving the Bellman equation with a non-constant investment opportunity set is not likely to change this, since by homotheticity it follows that J must be on the form given above.

This does not preclude the possibility that $\tilde{V}(W_t,\pi_t,t) = \hat{J}(W_t,t)$ for some function $\hat{J}$ for some other market structure than the simple one that we have chosen. Duffie and Epstein (1992a) is an example of this. Other examples include Ai (2012) and Bansal and Yaron (2004), who employ a richer economic environment than we do. The problem with this is that by changing too many features of the standard model, it may be hard to infer what really solved the problem, (in which case we have not really learnt all that much). In economics there is a long tradition with Ockham's razor.

7 Discussion of the results for the ordinally equivalent version

Returning to our earlier expression given in (32) for the risk premium of any risky asset having return rate $\mu_R(t)$ and volatility of return $\sigma_R(t)$, and the equilibrium interest rate given in (34), we can now formulate one main result:

Theorem 1 *In the ordinally equivalent model specified in Sections 3-6, there exists an equilibrium in which risk premiums of risky assets and the real interest rate are given by (41) and (42) respectively, where ρ is the time preference, γ the relative risk aversion, and β the impatience rate.*

As claimed risk premiums in the resulting model are linear combinations of the consumption-based CAPM and the market-based CAPM at each time $t \in [0,T]$.

In order to study the flexibility of the model, Table 3 illustrates additional calibrations to the ones presented in Section 2.2 for the recursive utility model of this section, consistent with the consumption and equity data summarized in Table 2. Here we consider the bills as risk free, and ignore the extra equation for the Government bills.

The "Kelly Criterion" means logarithmic utility in the standard model, which here corresponds to $\gamma = 1$ [8]. We notice from (41) that $\gamma = 1$ can not explain the observed risk premium. However, for values of γ close to 1 the model may give plausible results. If $\gamma = 1.05$ for example, we get the calibrated values $\rho = .97$ and $\beta = .04$. Thus a relative risk aversion of γ reasonably close to 1 can be consistent with data provided ρ is also close to 1 (from below), which is interesting and perhaps a bit surprising.

All the value sets presented in Table 3 represent exact fits to the consumption and stock market data summarized in Table 2. The CAPM++ version has acceptable values for risk aversion and the impatience rate, as we have seen before. By CAPM++ in Table 2 is meant the current version in continuous time, with an associated level of interest rate attached, and based on recursive utility. The original equilibrium model developed by Mossin (1966) was in a one period (timeless) setting with consumption only on the terminal time point, in which case wealth equals consumption. Since there was no consumption on the initial time point, no intertemporal aspects of consumption transfers arose in the classical model. This naturally corresponds to $u(c) = c$ for the the felicity index regarding consumption transfers, meaning $\rho = 0$ and $\psi = 1/\rho = +\infty$, and corresponding to perfect substitutability of consumption across time.

When the instantaneous correlation coefficient $\kappa_{Mc}(t)$ of returns and the aggregate consumption growth rate is small, our model handles this situation much better than the conventional one. The extreme case when $\kappa_{Mc}(t) = 0$ is, for example, consistent with the solution presented above for $\rho = 0$, which gives reasonable parameter values for the other parametes. The standard model then breaks down. If this correlation is small, the discrepancy between the standard model and the present one is even more striking than when $\hat{\kappa}_{Mc} = .4033$ (as it is for this set of data).

For example, in the hypothetical situation that $\kappa_{Mc}(t) = .01$, but the rest of the summary statistics are as in Table 2, the conventional model gives $\gamma = 1063$ and $\beta = 694$, while our model is consistent with $\beta = .027$, $\gamma = 1.54$, $\rho = .61$ corresponding to $EIS = 1.64$. The reason for this is that the second term in the equity premium is unaffected of $\kappa_{Mc}(t)$, and a decrease in $\kappa_{Mc}(t)$

[8] The value of $\gamma = 1$ is not in the permissible range for this parameter, and formally requires $h(v) = \ln(v)$.

	γ	ρ	EIS	β
Standard Model	26.37	26.37	.037	- .014
$\beta = .015$ fixed	.46	1.34	.74	.015
$\beta = .020$ fixed	.90	1.06	.94	.020
$\beta = .023$ fixed	1.15	.90	1.11	.023
$\beta = .030$ fixed	1.93	.41	2.43	.030
$\beta = .035$ fixed	2.14	.18	5.56	.035
$\rho = .90$ fixed	1.15	.90	1.11	.023
$\rho = .80$ fixed	1.30	.80	1.25	.025
$\rho = .50$ fixed	1.72	.50	2.00	.030
$\rho = .40$ fixed	1.86	.40	2.50	.031
CAPM ++	2.54	0.00	$+\infty$	.040
$\gamma = 0.50$ fixed	0.50	1.29	0.79	.011
$\gamma = 1.05$ fixed	1.05	.97	1.03	.040
$\gamma = 1.50$ fixed	1.50	.66	1.51	.027
$\gamma = 2.00$ fixed	2.00	.30	3.33	.033
$\gamma = 2.30$ fixed	2.30	.07	14.30	.036

Table 3: Various Calibrations Consistent with Table 2

only leads to a slight increase in the difference $(\gamma - \rho)$, while the expression for the interest rate is unaffected by $\kappa_{Mc}(t)$.

Most of the plausible calibration points for this model are located in the early resolution part of the (ρ, γ)-plane where $\gamma > \rho$. It is here that our results deviate from earlier research on recursive utility applied to explaining the historical equity premium and the interest rate, including the other empirical regularities.

The model can also produce late resolution solutions at otherwise low values of the parameters. As an example, if $\rho = 1.1$, this is consistent with $\beta = .02$ and $\gamma = .90$. The square root utility function is used in many examples in various textbooks; for $\gamma = .5$ the model calibrates to $\beta = .01$ and $\rho = 1.29$, i.e., late resolution but otherwise for reasonable values of the parameters.

7.1 Some new features of the model

It is reassuring that the risk premium of any risky asset depends on other investment opportunities in the financial market, and not just on this asset's covariance rate with consumption.

It is also satisfying that the return rate on government bonds depend on more than just the growth rate and the variance rate of aggregate consumption, but also on characteristics of other investment opportunities in the financial market.

Faced with increasing consumption uncertainty, the 'prudent' consumer will save and the interest rate accordingly falls in equilibrium. This is precautionary savings, and takes place in our model if $(1-\rho\gamma)/(1-\rho) > 0$, which then becomes the natural definition of prudence for this version of recursive utility. When the uncertainty of the return of the market portfolio increases, the recursive utility agent will buy bonds and sell stocks provided $\gamma > \rho$ and $\rho < 1$, or if $\gamma < \rho$ and $\rho > 1$, and will otherwise borrow and buy stocks.

This kind of analysis has no place in the conventional model, since there is no direct connection to the securities market in the expression for the equilibrium interest rate in (41), nor is there any direct connection to the securities market for the risk premium in (42).

8 The analysis for the nonordinal model

For this model we use the stochastic maximum principle of Section 4. The first order conditions are given by

$$\alpha\,\pi_t = Y_t\,\frac{\partial f}{\partial c}(c_t, V_t) \qquad \text{a.s. for all } t \in [0,T] \tag{43}$$

where $f = f_1$ is given in (12). If we normalize so that $\tilde{\sigma}_V(t) := \sigma_V(t)V_t$, the variable $Z(t)$ is now $Z_t = \sigma_V(t)$ and the future utility process V_t satisfies the following dynamics

$$dV_t = \Big(-\frac{\beta}{1-\rho}\frac{c_t^{1-\rho} - V_t^{1-\rho}}{V_t^{-\rho}} + \frac{1}{2}\gamma V_t \sigma_V'(t)\sigma_V(t)\Big)dt + V_t\sigma_V(t)dB_t \tag{44}$$

where $V(T) = 0$. This is the backward equation.

The function f of Section 4 is given by

$$f(t,c,v,\sigma_v) = f_1(c,v) - \frac{1}{2}A(v)v^2\sigma_v\sigma_v,$$

and since $A(v) = \gamma/v$, from (19) the adjoint variable Y has dynamics

$$dY_t = Y_t\big(\big\{f_v(c_t,V_t) + \frac{1}{2}\gamma\,\sigma_V(t)\sigma_V(t)\big\}\,dt - \gamma\sigma_V(t)\,dB_t\big), \tag{45}$$

where $Y(0) = 1$. From the FOC in (43) we get the dynamics of the state price deflator. We notice that Y is no longer a bounded variation process, and by Ito's lemma

$$d\pi_t = f_c(c_t,V_t)\,dY_t + Y_t\,df_c(c_t,V_t) + dY_t df_c(c_t,V_t). \tag{46}$$

By the adjoint and the backward equations this is

$$
\begin{aligned}
d\pi_t &= Y_t\, f_c(c_t, V_t)\, \Big(\{ f_v(c_t, V_t) + \frac{1}{2}\gamma \sigma_V'(t)\sigma_V(t) dt \} - \gamma \sigma_V(t) dB_t \Big) \\
&+ Y_t\, \frac{\partial f_c}{\partial c}(c_t, V_t)\, dc_t + Y_t\, \frac{\partial f_c}{\partial v}(c_t, V_t)\, dV_t + dY_t df_c(c_t, V_t) \\
&+ Y_t \Big(\frac{1}{2}\, \frac{\partial^2 f_c}{\partial c^2}(c_t, V_t)\, (dc_t)^2 + \frac{\partial^2 f_c}{\partial c\, \partial v}(c_t, V_t)\, (dc_t)(dV_t) + \frac{1}{2}\, \frac{\partial^2 f_c}{\partial v^2}(c_t, V_t)\, (dV_t)^2 \Big).
\end{aligned} \tag{47}
$$

Here

$$
f_c(c, v) := \frac{\partial f(c, v)}{\partial c} = \beta c^{-\rho} v^{\rho},\ f_v(c, v) := \frac{\partial f(c, v)}{\partial v} = -\frac{\beta}{1-\rho}(1 - \rho c^{1-\rho} v^{\rho-1}),
$$

$$
\frac{\partial f_c(c, v)}{\partial c} = -\beta \rho c^{-(1+\rho)} v^{\rho}, \quad \frac{\partial f_c(c, v)}{\partial v} = \beta \rho v^{\rho-1} c^{-\rho},
$$

$$
\frac{\partial^2 f_c}{\partial c^2}(c, v) = \beta \rho(\rho + 1) v^{\rho} c^{-(\rho+2)}, \quad \frac{\partial^2 f_c}{\partial c\, \partial v}(c, v) = -\beta \rho^2 v^{\rho-1} c^{-(\rho+1)},
$$

and

$$
\frac{\partial^2 f_c}{\partial v^2}(c, v) = \beta \rho(\rho - 1) v^{\rho-2} c^{-\rho}.
$$

8.1 The risk premiums

Denoting as before the dynamics of the state price deflator by

$$
d\pi_t = \mu_\pi(t)\, dt + \sigma_\pi(t)\, dB_t, \tag{48}
$$

from (47) and the above expressions we obtain the drift and the diffusion terms of π_t as

$$
\begin{aligned}
\mu_\pi(t) = \pi_t \big(&- \beta - \rho \mu_c(t) + \frac{1}{2}\rho(\rho + 1)\sigma_c'(t)\sigma_c(t) \\
&+ \rho(\gamma - \rho)\sigma_c'(t)\sigma_V(t) + \frac{1}{2}(\gamma - \rho)(1 - \rho)\sigma_V'(t)\sigma_V(t) \big)
\end{aligned} \tag{49}
$$

and

$$
\sigma_\pi(t) = -\pi_t \big(\rho \sigma_c(t) + (\gamma - \rho)\sigma_V(t) \big) \tag{50}
$$

respectively.

Recalling that the risk premium of any risky security with return process R is given by

$$
\mu_R(t) - r_t = -\frac{1}{\pi_t}\, \sigma_{R\pi}(t), \tag{51}
$$

it follows immediately from (50) and (51) that the formula for the risk premium of any risky security R is

$$\mu_R(t) - r_t = \rho\, \sigma_{Rc}(t) + (\gamma - \rho)\sigma_{RV}(t). \tag{52}$$

which is seen to be the same expression as found for the ordinally equivalent version in Section 5.1, given in (32) in terms of $\sigma_V(t)$. Next we turn to the equilibrium interest rate.

8.2 The equilibrium interest rate

The equilibrium short-term, real interest rate r_t is given by the formula

$$r_t = -\frac{\mu_\pi(t)}{\pi_t}, \tag{53}$$

as noticed in Section 5.2. In order to find an expression for r_t in terms of the primitives of the model, we use (49). We then obtain the following

$$r_t = \beta + \rho\mu_c(t) - \frac{1}{2}\rho(\rho+1)\sigma_c'(t)\sigma_c(t) - \\ \rho(\gamma-\rho)\sigma_{cV}(t) - \frac{1}{2}(\gamma-\rho)(1-\rho)\sigma_V'(t)\sigma_V(t). \tag{54}$$

Again, this is the same expression for the equilibrium interest rate as was obtained for the ordinally equivalent version based on (14), given in (34) in terms of $\sigma_V(t)$.

It follows from the discussion in Sections 5.1 and 5.2 that the potential for these fundamental relationships to solve the puzzles is the same for the present version of the recursive utility model.

We proceed to link the volatility term $\sigma_V(t)$ to an observable (or estimable) quantity in the market.

8.3 The determination of the volatility of future utility

In order to determine $\tilde{\sigma}_{V_t}$, i.e., to solve the adjoint equation, we proceed as follows: By the definition of directional derivatives we have that

$$\bigtriangledown U(c^*; c^*) = \lim_{\alpha\downarrow 0}\frac{U(c^* + \alpha c^*) - U(c^*)}{\alpha} = \lim_{\alpha\downarrow 0}\frac{U(c^*(1+\alpha)) - U(c^*)}{\alpha} \\ = \lim_{\alpha\downarrow 0}\frac{(1+\alpha)U(c^*) - U(c^*)}{\alpha} = \lim_{\alpha\downarrow 0}\frac{\alpha U(c^*)}{\alpha} = U(c^*),$$

where the third equality uses that U is homogeneous of degree one. It also follows from the linearity of the directional derivative that

$$\bigtriangledown U(c^*; c^*) = E\Big(\int_0^T \pi_t c_t^* \, dt \Big) = W_0 \pi_0$$

where W_0 is the wealth of the representative agent at time zero. Thus $U(c^*) = \pi_0 W_0$.

Let $V_t(c_t^*)$ denote future utility at the optimal consumption for our representation. Since this function is also homogeneous of degree one, the same, basic relationship holds here for the associated directional derivatives, i.e.,

$$\bigtriangledown V_t(c^*; c^*) = E_t\Big(\int_t^T \pi_s c_s^* \, ds \Big) = W_t \pi_t$$

and $\bigtriangledown V_t(c^*; c^*) = V_t(c^*)$. Together this shows that

$$V_t := \pi_t W_t \tag{55}$$

at the optimal consumption path c^*, so that for recursive utility the optimal future utility is a function of both wealth and the state price deflator π_t. Recall that π_t is not a Markov process. Since the wealth at any time t is given by

$$W_t = \frac{1}{\pi_t} E_t\Big(\int_t^T \pi_s c_s^* \, ds \Big),$$

and thus depends on past consumption and utility from time zero to the present time t, dynamic programming does not seem appropriate.

Since the Ito process V_t is a function of the agent's wealth and the state price deflator, it is a consequence of Ito's lemma that its diffusion term is

$$\tilde{\sigma}_V(t) = \sigma_\pi(t) W_t + \sigma_W(t) \pi_t. \tag{56}$$

As before we use (50) and (36), and observe that in equilibrium $\varphi_t' \cdot \sigma(t) = \sigma_M(t)$, so that by (36), $\sigma_W(t) = W_t \sigma_M(t)$. This gives

$$\tilde{\sigma}_V(t) = -\pi_t W_t(\rho \sigma_c(t) + (\gamma - \rho)\sigma_V(t)) + W_t \sigma_M(t)\, \pi_t.$$

Since by (44) $\tilde{\sigma}_V(t) = V_t \sigma_V(t)$, we get the following equation for $\sigma_V(t)$

$$\sigma_V(t) = \sigma_M(t) - \rho \sigma_c(t) - (\gamma - \rho)\sigma_V(t),$$

from which it follows that

$$\sigma_V(t) = \frac{1}{1 + \gamma - \rho} \Big(\sigma_M(t) - \rho \sigma_c(t) \Big). \tag{57}$$

By comparing with (40) for the ordinally equivalent model, at this point the two versions are seen to differ. As with the expected utility model, also recursive utility has reminiscences of cardinality. The unnormalized aggregator (f_1, A) is convenient for obtaining the desired disentangling by changing A with f fixed. Such a change in risk aversion is much less readily described in terms of the normalized aggregator $(f_2, 0)$.

8.4 The complete representation of the nonordinal version

In the expressions for the equilibrium risk premiums and the real interest rate $\sigma_V(t)$ was the only undetermined quantity. Inserting (57) into (52) and (54) we obtain the expressions

$$\mu_R(t) - r_t = \frac{\rho}{1+\gamma-\rho}\,\sigma_R(t)\sigma_c(t) + \frac{\gamma-\rho}{1+\gamma-\rho}\,\sigma_R(t)\sigma_M(t), \tag{58}$$

and

$$\begin{aligned} r_t &= \beta + \rho\mu_c(t) \\ &\quad - \Big(\frac{1}{2}\,\rho(1+\rho) - \frac{\rho^2(\gamma-\rho)}{1+\gamma-\rho} + \frac{1}{2}\frac{\rho^2(\gamma-\rho)(1-\rho)}{(1+\gamma-\rho)^2}\Big)\,\sigma_c'(t)\sigma_c(t) \\ &\qquad + \frac{\gamma\rho(\rho-\gamma)}{(1+\gamma-\rho)^2}\,\sigma_c'(t)\sigma_M(t) - \frac{1}{2}\frac{(\gamma-\rho)(1-\rho)}{(1+\gamma-\rho)^2}\,\sigma_M'(t)\sigma_M(t). \end{aligned} \tag{59}$$

It can be verified that dynamic programming, with our simple model for a financial market, does not lead to these expressions.

The main results in this section are then summarized as

Theorem 2 *For the nonordinal model specified in Sections 3, 4 and 8, there exists an equilibrium in which the equilibrium risk premium of any risky asset R is given by (58) and and the equilibrium real interest rate by (59). Here ρ is the time preference, γ the relative risk aversion, and β the impatience rate.*

Again the resulting model for risk premiums is a linear combinations of the consumption-based CAPM and the market-based CAPM at each time $t \in [0, T]$.

We now calibrate the model to the data summarized in Table 2. First we solve equation (58) and (59) together with the analogous equation for the Government bills. For a risk premium of Government bills equal to .50%, the unique calibration point for the data summarized in Table 2 is

$$\beta = 0.036, \qquad \gamma = .63, \qquad \rho = 2.14 \qquad (EIS = .47),$$

and is marked as Calibr 2 in Figure 1. The parameter values are all relatively small, so both the Equity Premium Puzzle as well as the Risk-Free Rate Puzzle seem less puzzling based on this model relative to both the standard model, as well as the recursive model based on dynamic programming.

However, the results of the ordinally equivalent model are in addition consistent with preference for early resolution of uncertainty. While preference for late resolution may be rational for many consumers, the issue of early resolution should play a major part in the solution of the puzzles.

We return to this issue in the Section 9, where we address heterogeneity among the agents.

Also the discrete time model (Aase (2013)) has two ordinally equivalent versions that match the data in a similar way to the two versions of this paper.

The results in the discrete time model are derived by directional derivatives, which give analogous results to those we get by the stochastic maximum principle. For example are the coefficients in terms of the preference parameters in the asoicated risk premiums exactly the same for both versions. In the discrete time model it is documented that these results are not the same as when dynamic programming is used.

8.5 Discussion of the results for the nonordinal model

In Table 4 we illustrate additional parameter values to the one presented above for the recursive utility model consistent with the data behind Table 2. The model appears to be very robust over a wide range of parameter values, as can be seen from this table.

It seems fair to say that the calibration points correspond to plausible values of the various parameters, but as noticed they are in the late resolution part of the (ρ, γ)-plane where $\gamma < \rho$. The main difference between the two ordinally equivalent versions is that the present model calibrates to values in the range $\gamma < \rho$, and $\rho > 1$, while the ordinally equivalent version calibrates to $\gamma > 1 > \rho$ and accordingly $EIS > 1$, for the data summarized in Table 2. Moreover $\rho = 0$ does not make sense for the nonordinal model, while this gives a reasonable calibration point for the ordinally equivalent version. However, the ordinally equivalent version is also consistent with calibrations in the region $0 < \gamma < \rho < 1$, corresponding to late resolution.

According to the present version of recursive utility, we may interpret the above calibration to mean that the typical consumer represented by the data in Table 2 is not all that active in the stock market. It is probably true that many people use the securities market for savings and long term

	γ	ρ	EIS	β
Model (58)-(59):				
$\beta = .010$ fixed	.002	1.62	.78	.010
$\beta = .020$ fixed	.345	1.94	.51	.020
$\beta = .030$ fixed	.634	2.21	.45	.030
$\beta = .040$ fixed	.877	2.44	.41	.040
$\beta = .060$ fixed	1.26	2.80	.36	.060
$\beta = .080$ fixed	1.56	3.08	.33	.080
$\beta = .100$ fixed	1.80	3.31	.30	.100
$\beta = .350$ fixed	3.25	4.67	.21	.350
$\gamma = 0.00$ fixed	0.00	1.62	.62	.01
$\rho = 1.00$ fixed	-0.66	1.00	1.00	.04

Table 4: Various Calibrations Consistent with Table 2.

investments, without worrying about daily or more frequent trade. More on this in Section 9.

We may then ask the hypothetical question: How large must the volatility of the market portfolio be in order for the representative agent to be concerned about early resolution? It can be seen from the expressions for the equity premium and the interest rate that when the volatility of the market portfolio increases, this will tend change the calibration points to the region where $\gamma > \rho$.

The critical value of $\sigma_M(t)$ for this data set is about .245. For example, for a value of $\sigma_M(t)$ corresponding to 27.4%, we get preference for early resolution of uncertainty, ceteris paribus, for the values $\beta = .00$, $\gamma = 4.48$, and $\rho = .64$.

For a market volatility of 28.3% the model is consistent with $\beta = .01$, $\gamma = 3.17$, and $\rho = .24$, for $\sigma_M(t) = .292$ the model is consistent with $\beta = .01$, $\gamma = 2.71$, and $\rho = .27$, e.t.c.

In some countries, this order of magnitude for the volatility of the market portfolio is not uncommon. An example is Norway where this volatility has been estimated to 35% for the period 1971-2012. When the market volatility is high, normally the risk premium is also high. Consider the following example.

Example 1. The data for Norway for the period 1971-2012[9] are: In real terms $\sigma_M(t) = 0.35$, risk premium = 8.71% and the risk-free rate is 2.25%. Assuming the rest of the summary statistic in real terms for Norway for the period indicated ($\sigma_{M,c}(t) = .00048$, $\sigma_c(t) = .024$, $\mu_c(t) = .0311$, $\mu_M(t) = .11$),

[9]Data made available by Thore Johnsen, based on http://www.ssb.no/statistikkbanken

this gives the calibrated values for the nonordinal model of this section

$$\beta = .01, \quad \gamma = 3.03 \quad \text{and} \quad \rho = .58 \quad (EIS = 1.72)$$

i.e., preference for early resolution. By fixing $\rho = .8$, we get

$$\beta = .001 \quad \text{and} \quad \gamma = 3.25 \quad (EIS = 1.25).$$

For the US-economy this value of ρ calibrates to negative values of β, and γ for this model. For the ordinally equivalent model we find the following calibration:

$$\beta = .01 \quad \gamma = 1.21 \quad \text{and} \quad \rho = 1.73$$

i.e., preference for late resolution at low parameter values. (There is also a calibration point $\gamma = 1.01$, and $\rho = 1.02$ for this value of β.)

This magnitude of the volatility of the market portfolio is common in many countries, but Norway has the highest stock market volatility in the following group: France, Germany, The Netherlands, Sweden, Switzerland, UK, Japan and US. $\square$

This example demonstrates that the representative agent may be of the type of this section, and calibrate to data consistent with preference for early resolution.

In the real market places there may be people favoring early resolution of uncertainty, and others preferring late. We therefore consider a heterogeneous model with agents of the two different kinds of recursive utilities studied in this paper, which is the topic of our last section.

9 Heterogeneity in preferences

From the above results, it seems reasonable to study a model with heterogeneity containing two agents, one of each kind of recursive utility studied in this paper. Since each agent fits the data in isolation with reasonably low parameter values, it would seem likely that so would a model with a representative agent. This would allow us to present an economy consisting of two groups of people, one more concerned with stock market uncertainty than the other. This is what we formalize next.

9.1 The Arrow-Debreu economy

In this section we derive an Arrow-Debreu markets equilibrium in which each agent has a recursive utility function U_i of the type we have considered in this paper. As in Duffie (1986) there exists an implementation of such equilibria is the setting with security and spot markets only, given an appropriate set of admissible trading strategies and a spanning assumption on nominal cumulative dividend processes.

As we shall simply calculate the relevant equilibrium, we do not really employ the theorems for such equilibria to exist, but there is a theory for recursive preferences in this regard that should be consulted (Duffie, Geoffard and Skiadas (1994)).

The situation is as follows: Given an initial allocation $(e^1, e^2, \ldots, e^m) \in L^m$, an m-dimensional Ito process, with $e = \sum_i e^i$, an *equilibrium* is a feasible allocation $(c^1, c^2, \ldots, c^m)$ and a non-zero linear price functional $\Pi : L \to R$ such that, for all i, c^i solves the problem

$$\max_{c \in L} U_i(c) \quad \text{subject to} \quad \Pi(c) \leq \Pi(e^i) \tag{60}$$

By assuming there is no arbitrage possibilities in this market of Arrow-Debreu securities, the price functional is strictly positive on L, hence it is bounded, and thus also continuous. By the Riesz' Representation Theorem there is an element $\pi \in L$, the Riesz Representation, such that $\Pi(c) = E\big(\int_0^T \pi_t c_t\big)$ for any $c \in L$.

Under certain smoothness conditions on the aggregator (f, A), there exists an Arrow-Debreu equilibrium $(\Pi, (c^1, c^2, \ldots, c^m))$ having the following properties:

(i) $(c^1, c^2, \ldots, c^m)$ is Pareto optimal. (ii) For each i, U_i has a gradient at c^i with a Riesz Representation $\pi_i(c^i)$ given by the stochastic maximum principle in (21) for our models ($m = 2$), or

$$\alpha_i \pi_t^i = Y_t^i \, \frac{\partial f^i}{\partial c}(c_t^i, V_t^i) \quad \text{a.s. for all } t \in [0, T], \text{ i} = 1, 2. \tag{61}$$

(iii) The state price deflator $\pi_t = \alpha_i \pi_t^i$ for some constants $\alpha_i > 0$, $i = 1, 2$.

The condition (iii) can be considered as a version of Borch's characterization of Pareto optimality (Borch (1960-62)).

Equality in the budget constraints determine the constants α_i as a function of the preferences of the agents and the joint probability distributions of the initial endowments. Defining the agent weights $\lambda_i = \alpha_i^{-1}$ for $i = 1, 2, \ldots, m$, the function $U_\lambda(c) = \sum_{i=1}^m \lambda_i \, U_i(c^i)$ can be thought of as the utility function of the representative agent, here as a generalized recursive utility function, where $c_t = \sum_{i=1}^m c_t^i = e_t$. In the following section we consider the case of $m = 2$.

9.2 Heterogeneity with U_1 and U_2

Inspired by the calibrations of the previous sections, we imagine that the market consists of two groups of people, one with recursive preferences belonging to the ordinally equivalent specification, the other group having preferences belonging to the nonordinal version, and want to characterize the resulting Pareto optimal equilibrium.

The following results are proven in Appendix 1. We denote $c_t^{(i)}/c_t$ as the optimal fraction of the aggregate consumption consumed by agent i at time t, $i = 1, 2$.

9.3 The risk premium with heterogeneity

The risk premium of a risky asset denoted R has the following representation

$$\mu_R(t) - r_t = \frac{1}{\bar{\psi}_t}\Big(\sigma_c(t)\sigma_R(t) + \big(\frac{c_t^{(1)}}{c_t}\big)\frac{\gamma_1 - \rho_1}{\rho_1}\,\sigma_{V_1}(t)\sigma_R(t) \\ + \big(\frac{c_t^{(2)}}{c_t}\big)\frac{\gamma_2 - \rho_2}{\rho_2}\,\sigma_{V_2}(t)\sigma_R(t)\Big), \quad (62)$$

where the average value of the population EIS ($\psi_i = 1/\rho_i$) is

$$\bar{\psi}_t := \frac{1}{\rho_1}\Big(\frac{c_t^{(1)}}{c_t}\Big) + \frac{1}{\rho_2}\Big(\frac{c_t^{(2)}}{c_t}\Big). \quad (63)$$

and the average time preference is

$$\bar{\rho}_t := \rho_1\big(\frac{c_t^{(1)}}{c_t}\big) + \rho_2\big(\frac{c_t^{(2)}}{c_t}\big).$$

Since the harmonic mean is smaller than or equal to the arithmetic mean, note that $1/\bar{\psi}_t \leq \bar{\rho}_t$. Also

$$\sigma_{V_1}(t) = \frac{\sigma_M(t) - \frac{1}{\bar{\psi}_t}\,\sigma_c(t)}{1 + \frac{1}{\bar{\psi}_t}\big(\frac{\gamma_1-\rho_1}{\rho_1}\big(\frac{c_t^{(1)}}{c_t}\big) + \frac{\gamma_2-\rho_2}{(1-\gamma_2)\rho_2}\big(\frac{c_t^{(2)}}{c_t}\big)\big)}. \quad (64)$$

and

$$\sigma_{V_2}(t) = \frac{\sigma_M(t) - \frac{1}{\bar{\psi}_t}\,\sigma_c(t)}{1 - \gamma_2 + \frac{1}{\bar{\psi}_t}\big(\frac{(\gamma_1-\rho_1)(1-\gamma_2)}{\rho_1}\big(\frac{c_t^{(1)}}{c_t}\big) + \frac{\gamma_2-\rho_2}{\rho_2}\big(\frac{c_t^{(2)}}{c_t}\big)\big)}. \quad (65)$$

Remark 1: As a consistency check, notice that when $c_t^{(1)} = c_t$ for all t a.s., then (62) reduces to the risk premium of model 1 in Section 4, the nonordinal version, and when $c_t^{(2)} = c_t$, for all t a.s., then (62) reduces to the corresponding risk premium of model 2 in Section 4, the ordinally equivalent one. When $\gamma_1 = \rho_1$ and $\gamma_2 = \rho_2$ (62) reduces to the risk premium of the standard model with two heterogeneous agents. □

Remark 2: Note that when $\sigma_{V_1}(t) \equiv 0$ it is not the case that we get the risk premiums of the model 2, and vice versa. This stems from the market clearing condition for the representative agent, and both $\sigma_{V_1}(t)$ and $\sigma_{V_2}(t)$ depend on all the preference parameters in the model (except the β's) and in a non-symmetrical way. □

9.4 The equilibrium interest rate with heterogeneity

The equilibrium short rate for the heterogeneous model is given by the following expression

$$
\begin{aligned}
r_t = {} & \bar{\beta}_t^{(\rho)} + \frac{1}{\bar{\psi}_t}\,\mu_c(t) \\
& -\frac{1}{2}\frac{1}{\bar{\psi}_t}\Big\{\big(\frac{c_t^{(1)}}{c_t}\big)\,\frac{1}{\rho_1}\big(\frac{1}{\rho_1}+1\big)\,\frac{1}{\bar{\psi}_t^2}\Big(\sigma_c(t)+\big(\frac{c_t^{(1)}}{c_t}\big)\frac{\gamma_1-\rho_1}{\rho_1}\,\sigma_{V_1}(t)+\big(\frac{c_t^{(2)}}{c_t}\big)\frac{\gamma_2-\rho_2}{\rho_2}\,\sigma_{V_2}(t)\Big)^2 \\
& +\big(\frac{c_t^{(2)}}{c_t}\big)\,\frac{1}{\rho_2}\big(\frac{1}{\rho_2}+1\big)\,\frac{1}{\bar{\psi}_t^2}\Big(\sigma_c(t)+\big(\frac{c_t^{(1)}}{c_t}\big)\frac{\gamma_1-\rho_1}{\rho_1}\,\sigma_{V_1}(t)+\big(\frac{c_t^{(2)}}{c_t}\big)\frac{\gamma_2-\rho_2}{\rho_2}\,\sigma_{V_2}(t)\Big)^2\Big\} \\
& +\frac{1}{\bar{\psi}_t^2}\Big\{\big(\frac{c_t^{(1)}}{c_t}\big)\,\frac{\rho_1-\gamma_1}{\rho_1^2}\Big(\sigma_c(t)+\big(\frac{c_t^{(1)}}{c_t}\big)\frac{\gamma_1-\rho_1}{\rho_1}\,\sigma_{V_1}(t)+\big(\frac{c_t^{(2)}}{c_t}\big)\frac{\gamma_2-\rho_2}{\rho_2}\,\sigma_{V_2}(t)\Big)\sigma_{V_1}(t) \\
& +\big(\frac{c_t^{(2)}}{c_t}\big)\,\frac{\rho_2-\gamma_1}{\rho_2^2}\Big(\sigma_c(t)+\big(\frac{c_t^{(1)}}{c_t}\big)\frac{\gamma_1-\rho_1}{\rho_1}\,\sigma_{V_1}(t)+\big(\frac{c_t^{(2)}}{c_t}\big)\frac{\gamma_2-\rho_2}{\rho_2}\,\sigma_{V_2}(t)\Big)\sigma_{V_2}(t)\Big\} \\
& -\frac{1}{2}\frac{1}{\bar{\psi}_t}\Big\{\big(\frac{c_t^{(1)}}{c_t}\big)\,\frac{\rho_1-\gamma_1}{\rho_1}\,\gamma_1\,\frac{\rho_1-1}{\rho_1}\,\sigma_{V_1}(t)^2+\big(\frac{c_t^{(2)}}{c_t}\big)\,\frac{\rho_2-\gamma_2}{\rho_2}\,\gamma_2\,\frac{\rho_2-1}{\rho_2}\,\sigma_{V_2}(t)^2\Big\}
\end{aligned}
\tag{66}
$$

where

$$\bar{\beta}_t^{(\rho)} := \frac{\sum_{i=1}^{2}\big(\frac{c_t^{(i)}}{c_t}\big)\big(\frac{\beta_i}{\rho_i}\big)}{\bar{\psi}_t}. \tag{67}$$

Note that when $\beta_1 = \beta_2 := \beta$, then $\bar{\beta}_t^{(\rho)} = \beta$ for all t.

Remark 3: As with the risk premium, one can see that when $c_t^{(1)} \equiv c_t$, then (66) reduces to the interest rate in model 1 of Section 4, the nonordinal version, and when $c_t^{(2)} \equiv c_t$, then (66) reduces to the corresponding interest rate in model 2 of Section 4, the ordinally equivalent one. When $\gamma_1 = \rho_1$ and $\gamma_2 = \rho_2$, then (66) reduces to the equilibrium interest rate in the standard model with two heterogeneous agents. $\square$

Remark 4: Consistent with the results in Sections 8.1-2, at this level the interest rate looks symmetric in the two agents, but market clearing for the representative agent implies that this symmetry does not carry over to the two volatilities $\sigma_{V_1}(t)$ and $\sigma_{V_2}(t)$. $\square$

We summarize our results as

Theorem 3 *For the heterogeneous model there exists an equilibrium in which the equilibrium risk premium of any risky asset R is given by (62) and and the equilibrium real interest rate by (66). Here $\bar{\rho}_t$ is time preference, $\bar{\psi}_t$ the EIS, and $\bar{\beta}_t^{(\rho)}$ the impatience rate of the representative agent, where $1/\bar{\psi}_t \leq \bar{\rho}_t$.*

9.5 Applications of the heterogeneous model

To demonstrate these results, assume that agent 1 consumes 11/12 of the total consumption, and consider a situation where ρ_1 is to be determined together with $\bar{\beta}_t^{(\rho)}$ such that (62) matches the estimated equity premium of 5.98% and (66) matches the short rate of .0080 for the period considered, together with the rest of the summary statistics of Table 2, when the other parameters are as follows: $\gamma_2 = 2.3$, $\rho_2 = .60$ and $\gamma_1 = 1.8$. The result is $\rho_1 = 3.70$ and $\bar{\beta}_t^{(\rho)} = .033$.

This implies preference for early resolution of uncertainty for agent 2, and late for agent 1. According to Table 4 agent 1 would in isolation calibrate to the same data with a value of $\rho_1 = 3.31$ and an impatience rate $\beta_1 = .10$. Increasing the parameter γ_1, increases ρ_1 and decreases the parameter $\bar{\beta}_t^{(\rho)}$. For example, when $\gamma_1 = 2.3$ we obtain $\rho_1 = 4.0$ and $\bar{\beta}_t^{(\rho)} = .025$. In this situation $\sigma_{V_1,M}(t) = -.12$, and $\sigma_{V_2,M}(t) = .12$.

Let us think of agent 1 representing "the public at large". Since agent 1 dominates in numbers (= aggregate consumption), an interpretation of the above is that when these two groups represent the market together, the market as a whole will prefer late resolution of uncertainty, at about the same impatience rate as if agent 1 alone represents the market. This group, less concerned with the stock market uncertainty, has more problems with consumption substitution ($\psi_1 < 1$) than the other group ($\psi_2 > 1$). This means that at least this group could benefit from the existence of a well functioning pension insurance market. In real life individuals with a finite life span can in general not time diversity, so the people in group 2 will also presumably benefit from such a market, provided, for example, the insurance market takes the role of time diversification seriously. There should in principle be no problems with this, since a stable, long term risk premium of around 6% is consistent with the model. Agent 1 is not much risk averse in this example ($\gamma_1 = 1.8 < \gamma_2 = 2.3$). In real life the means to invest typically resides with the agent 2. With a low interest rate, ordinary people may then be inclined to borrow, given that credit is provided.

The preference parameters of the representative agent may, for this illustration, be summarized as

$$\bar{\rho}_t = 3.26, \quad \bar{\psi}_t = .49, \quad \bar{\gamma}_t = 1.84 \quad \text{and} \quad \bar{\beta}_t^{(\rho)} = .039,$$

where $\rho_1 = 3.50$, $\gamma_1 = 1.8$, $\beta_1 = .039$ and $\rho_2 = .60$, $\gamma_2 = 2.3$, and $\beta_2 = .039$. Note that $1/\bar{\psi}_t = 2.04 < \bar{\rho}_t = 3.26$. Group 2 manages consumption substitution well ($\psi_2 > 1$), and early resolution holds for this group ($\gamma_2 > \rho_2$).

Another calibration is: $\rho_1 = 3.58$, $\gamma_1 = 2.0$, $\beta_1 = .042$ and $\rho_2 = .60$, $\gamma_2 = 1.76$, and $\beta_2 = 0.042$. In this case $\sigma_{V_1,M}(t) = -.035$, and $\sigma_{V_2,M}(t) = .046$. For asset allocation, the latter volatilities are reasonable.

Naturally, if group 2 represents the majority in the market, results would look different. If this is the case, also the representative agent is likely to prefer early resolution of uncertainty, depending on the composition of the population.

In conclusion, a variety of scenarios are possible, and the above illustrations may seem like plausible descriptions of an economy with statistics summarized in Table 2 for the period considered, consistent with the data. The advantage with our heterogeneous model is that the representative agent does not point to a monolithic type of preference, but gives room for a more realistic composition of the population.

9.6 Limited stock market participation

It is known that in the economy only a certain fraction of the population owns stock. As an illustration, suppose that this fraction is about 8-9% (see e.g., Vissing-Jørgensen (1999)). We suggest to use our heterogeneous model to see if the resulting model still explains the data with this added limitation. Let agent 2 represent the fraction that participates in the stock market, and agent 1 the non-participating fraction. We propose to capture this by setting to zero agent 1's part of future utility volatility that depends on the stock market. This means that in the expression for $\sigma_{V_1,M}(t)$, which is given by

$$\sigma_{V_1,M}(t) = \frac{\sigma_M(t)\sigma_M(t) - \frac{1}{\bar{\psi}_t}\,\sigma_c(t)\sigma_M(t)}{1 + \frac{1}{\bar{\psi}_t}\left(\frac{\gamma_1-\rho_1}{\rho_1}\left(\frac{c_t^{(1)}}{c_t}\right) + \frac{\gamma_2-\rho_2}{(1-\gamma_2)\rho_2}\left(\frac{c_t^{(2)}}{c_t}\right)\right)}$$

the part $\sigma_{M,M}(t) = 0$ for all t. This means that in the market clearing, only agent 2 holds the market portfolio.

Whether the second term in the numerator should also be changed can be questioned, since after all the agent 1 is still consuming (e.g., Vissing-Jørgensen (1999)). One view is that the consumption growth of non-stockholders covaries with the stock return in the same way as the consumption growth of stockholders. There are also arguments why consumption growth of non-stockholders is less correlated with stock returns than that of stockholders. In the present setting this argument becomes partly irrelevant, since there is no index $i = 1$ on the consumption correlation $\sigma_c(t)\sigma_M(t)$ in this expression.

The question is here if the future utility growth of non-stockholders covaries with the stock return in the same way as the future utility growth of stockholders. In our model these covariance rates are not equal to start with, as can be seen when comparing (64) to (65).

With the parameters similar to the above illustration, the model fits the data for a range of values for the relative risk aversion γ_1: Consider a situation where $\gamma_1 = 1.2$ and $\rho_2 = .9$. When $\gamma_2 = 2.8$ then $\rho_1 = 2.2$ and $\bar{\beta}_t^{(\rho)} = .095$. When γ_2 increases, ρ_1 stays the approximately same, and $\bar{\beta}_t^{(\rho)}$ decreases. For example, when $\gamma_2 = 3.2$, then $\bar{\beta}_t^{(\rho)} = .06$, and when $\gamma_2 = 3.7$, then $\bar{\beta}_t^{(\rho)} = .024$. In this situation, when $\gamma_2 = 3.7$, then $\sigma_{V_1,M}(t)$, and $\sigma_{V_2,M}(t)$ are too large to be reasonable.

The basic conclusions from the illustration in the last section remain. The major changes are that the the stockholders are more risk averse if they alone have to clear the financial market, and the time preference of agent 1 now appears smaller. The representative agent's impatience rate varies from 1.7% to about 10% when γ_2 varies from 3.8 to 2.7.

An alternative set of parameters is $\gamma_2 = 1.76$ and $\rho_2 = .60$, $\gamma_1 = 2.00$ and $\rho_1 = 2.93$, with $\bar{\beta}_t^{(\rho)} = .21$. In this case $\sigma_{V_1,M}(t) = -0.18$, and $\sigma_{V_2,M}(t) = 0.23$. Here the volatilities are more reasonable, while the impatience rate seems high.

According to Vissing-Jørgensen (1999), the postwar participation ratio has been higher than 8%, more of the order of 20%. Since parts of the data in Table 2 is postwar, we also check for this modification. Some results of the calibration are then: $\rho_1 = 2.87$, $\gamma_1 = 1.2$, $\rho_2 = .80$ and $\gamma_2 = 2.8$. The time impatience parameter $\bar{\beta}_t^{(\rho)}$ stays around 10%, when γ_2 varies between 2.8 and 3.8, with the rest of the parameters essentially as above. In this example the representative agent has parameters

$$\bar{\rho}_t = 2.46, \quad \bar{\psi}_t = .53, \quad \bar{\gamma}_t = 1.6 \quad \text{and} \quad \bar{\beta}_t^{(\rho)} = .10,$$

The associated impatience rate of the representative agent could result from about 15% in group 1, and about 4.4% in group 2. Here $\sigma_{V_1,M}(t) = .010$, and $\sigma_{V_2,M}(t) = -.012$.

Except from the impatience rate, and the change in the covariance rates of the future utilities with the market, this scenario is in many ways similar to the one with 8% participation[10], and the basic conclusions seem to remain.

10 Optimal asset allocation with recursive utility

From the results is Section 6.1, 8.3 and Appendix 1, the following results are immediate:

Corollary 1 *The optimal portfolio fractions for the nonordinal model are given by*

$$\varphi_1(t) = (\sigma_t \sigma_t')^{-1}\nu_t + \sigma_{V_1}(t)\sigma_t^{-1},$$

for the ordinally equivalent model by

$$\varphi_2(t) = (\sigma_t \sigma_t')^{-1}\nu_t + (1-\gamma_2)\sigma_{V_2}(t)\sigma_t^{-1},$$

and for the heterogeneous model by

$$\varphi(t) = (\sigma_t \sigma_t')^{-1}\nu_t + \big(\frac{c_t^{(1)}}{c_t}\big)\sigma_{V_1}(t)\sigma_t^{-1} + \big(\frac{c_t^{(2)}}{c_t}\big)(1-\gamma_2)\sigma_{V_2}(t)\sigma_t^{-1},$$

assuming the matrix σ_t is invertible.

Proof: Use that $\sigma_M(t) = \varphi'(t)\sigma(t)$. $\square$

Notice that the first term in these fractions is the one of the standard model having $\gamma = 1$. The optimal factions with recursive uility depend on both risk aversion and time preference through the terms $\sigma_{V_i}(t)$.

As an illustration, consider the standard situation with one risky and one risk-free asset for the following parameter values in the heterogeneous model: $\gamma_1 = 2.00$, $\rho_1 = 2.77$, $\gamma_2 = 1.76$ and $\rho_2 = .60$. For the data of Table 2 ($\sigma_M(t) = .16$, $\sigma_c(t) = .0355$), we obtain the following fractions: $\varphi_1 = .069$, $\varphi_2 = .80$ and $\varphi = .13$. Here $\sigma_{V_1}(t) = -.3633$ and $\sigma_{V_2}(t) = .3234$. This seems

[10] In this connection it can be of some interest to note that Andersen et. al. (2008) use controlled experiments with field subjects in Denmark to elicit the impatience rate and risk preference, ignoring the subject of time preferences. First, an estimate of β around 25% is reached assuming risk neutrality, second, a new estimate of β around 10% is obtained assuming risk aversion, with an associated estimate of γ around .74, both based on arithmetic averaging.

like a fair description of empirical regularities, since the typical household is stipulated to hold between 6% and 20% in equities. Instead of assuming that people in group 1 do not hold stock, with these parameter values they hold about 7% on average in equities, while people in the other group hold about 80% on average, in which case the "typical hosehold" holds about 13% in equities, on average. Again people in group 1 consume about 91-92% of total consumption in this illustration[11].

11 Extensions

There are many important issues to explore, based on the framework of this paper. The life cycle model, for instance, can be better understood once time preference is separated from risk aversion. This gives new insights in the comparisons of defined benefit to defined contribution pension plans. Starting with the life cycle model, and using market clearing, we have derived results identical to the ones in this paper, but from a very different starting point. This shows that our results are robust. We also have an application to the economics of climate change. Other applications are plentiful.

12 Conclusions

We have addressed the well-known empirical deficiencies of the conventional asset pricing model in financial and macro economics. It is known that the standard model with a separable and additive utility representation does not necessarily work all that well for *temporary* problems. In particular the standard model does not fit real data for plausible parameter values of the utility function. Our approach is to change only one feature with the conventional model, the preference structure. This solves both puzzles. It is the property with low values of the parameters in the preference relation that is new.

Recursive preferences deviate from the separable time additive case in several important ways, and it is not at all clear that standard methods work the way we are used to.

We use a general method of optimization, the stochastic maximum principle, which give the same results as directional derivatives in function space. For the standard model the results of this procedure coincides with those of dynamic programming. As this paper clearly demonstrates, this is not so

[11]In contrast, the standard one-agent model prescribes 117% in equities with $\gamma = 2$ for the average household. In order to get to a level of 13%, γ has to be of the order of 18.

with recursive utility. The stochastic maximum principle does not require the Markov property, and neither does the utility gradient method. Our approach is thus more generally applicable, and this supports our results.

An important limitation with the conventional model is the equality between risk aversion and time preference. With recursive utility properly defined, these inherently different properties of an individual are partly disentangled.

We present two ordinally equivalent versions of recursive utility. For the US-data one version calibrates to preference for early resolution of uncertainty, the other to late. It is demonstrated that for other sets of data this may be different. We suggest a heterogeneous model where the representative agent is derived from the two ordinally equivalent versions. The resulting model is found to explain well both the Equity Premium Puzzle as well as the Risk-Free Rate Puzzle.

Our models, both the one in continuous time as well as the discrete time version, are able to fit the high estimate of the equity premium of 6.18% related to the return on the S&P-500 index in the USA for the period of 1889-1978, the low estimate of .8% for the risk free real interest rate, the low estimate of the consumption volatility for the same period, the high volatility estimate of the market equity index, and the low estimate of the covariance between returns on equity and the growth rate of aggregate consumption, for reasonable parameter values of the utility function of the representative agent. The models also explain the covariances of Government bills with consumption and equity, for a moderate risk premium for the bills. Here the standard model can offer no reasonable answers at all. Our model with heterogeneity yields an explanation of the limited market participation issue, as well as the problem of optimal asset allocation.

One important aspect left out of this paper is about incomplete markets. Our findings are likely to have broad economic implications.

13 Appendix 1

1. THE RISK PREMIUMS AND THE SHORT RATE FOR THE HETEROGENEOUS MODEL.

In this section we prove the results for the risk premiums and the short rate for heterogeneous model. The methods used here are somewhat different from the ones used in the rest of the paper, since we have to find the optimal consumption for both agents separately. To this end, consider the expression for the aggregate consumption $c_t = c_t^{(1)} + c_t^{(2)}$, where the optimal consumptions for each of the agents are given in

$$c_t = (\beta_1\, Y_t^{(1)})^{\frac{1}{\rho_1}}\, V_1(t)\, \pi_t^{-\frac{1}{\rho_1}} + (\beta_2\, Y_t^{(2)})^{\frac{1}{\rho_2}}\, ((1-\gamma_2)V_2(t))^{\frac{\rho_2-\gamma_2}{\rho_2(1-\gamma_2)}}\, \pi_t^{-\frac{1}{\rho_2}}. \quad (68)$$

This follows from the first order conditions of optimal consumption in (61).

We now use Ito's lemma to obtain the dynamic equation for the aggregate consumption, taking into account that $Y_t^{(2)}$ is of bounded variation on compacts, but $Y_t^{(1)}$ is not. This gives

$$\begin{aligned}
dc_t = c_t\mu_c(t)dt + c_t\sigma_c(t)dB_t = \\
d\big((\beta_1 Y_t^{(1)})^{\frac{1}{\rho_1}}\big)V_1(t)\pi_t^{-\frac{1}{\rho_1}} + \big(\beta_1 Y_t^{(1)}\big)^{\frac{1}{\rho_1}} d(V_1(t))\pi_t^{-\frac{1}{\rho_1}} \\
+\big(\beta_1 Y_t^{(1)}\big)^{\frac{1}{\rho_1}} V_1(t) d(\pi_t^{-\frac{1}{\rho_1}}) + \big(\beta_1 Y_t^{(1)}\big)^{\frac{1}{\rho_1}} d(V_1(t)\pi_t^{-\frac{1}{\rho_1}}) \\
+\pi_t^{-\frac{1}{\rho_1}} d\big((\beta_1 Y_t^{(1)})^{\frac{1}{\rho_1}}(V_1(t)\big) + (V_1(t)) d\big(\beta_1 Y_t^{(1)}\big)^{\frac{1}{\rho_1}} \pi_t^{-\frac{1}{\rho_1}}\big) \\
+d\big((\beta_2 Y_t^{(2)})^{\frac{1}{\rho_2}}\big)\,\big((1-\gamma_2)V_2(t)\big)^{\frac{\rho_2-\gamma_2}{\rho_2(1-\gamma_2)}}\pi_t^{-\frac{1}{\rho_2}} \\
+\big(\beta_2 Y_t^{(2)}\big)^{\frac{1}{\rho_2}}\pi_t^{-\frac{1}{\rho_2}} d\big(\big((1-\gamma_2)V_2(t)\big)^{\frac{\rho_2-\gamma_2}{\rho_2(1-\gamma_2)}}\big) \\
+\big(\beta_2 Y_t^{(2)}\big)^{\frac{1}{\rho_2}}\big((1-\gamma_2)\,V_2(t)\big)^{\frac{\rho_2-\gamma_2}{\rho_2(1-\gamma_2)}}\, d\big(\pi_t^{-\frac{1}{\rho_2}}\big) \\
+(\beta_2 Y_t^{(2)})^{\frac{1}{\rho_2}}\, d\big(\big((1-\gamma_2)V_2(t)\big)^{\frac{\rho_2-\gamma_2}{\rho_2(1-\gamma_2)}}\pi_t^{-\frac{1}{\rho_2}}\big).
\end{aligned}$$

Also, by Ito's lemma

$$\begin{aligned}
d\pi_t^{-\frac{1}{\rho_i}} = \Big(-\frac{1}{\rho_i}\pi_t^{-\frac{1}{\rho_i}-1}\mu_\pi(t) + \frac{1}{2}\frac{1}{\rho_i}(\frac{1}{\rho_i}+1)\pi_t^{-(\frac{1}{\rho_i}+2)}\sigma_\pi^2(t)\Big)dt \\
-\frac{1}{\rho_i}\pi_t^{-\frac{1}{\rho_i}-1}\sigma_\pi(t)\, dB_t. \quad (69)
\end{aligned}$$

We now use the dynamics of the utility processes $V_i(t)$ and the adjoint processes Y_t^i given in Sections 5 and 8 together with the dynamics for $\pi_t^{-\frac{1}{\rho_i}}$ given in (69). This results in a stochastic differential equation for c, where the drift consists of 14 terms and the diffusion of 5. If we use the expressions given above for the optimal consumptions of the two agents $c_t^{(1)}$ and $c_t^{(2)}$ respectively, this reduces to the following

$$dc_t = c_t\mu_c(t)dt + c_t\sigma_c(t)dB_t =$$

$$\big\{c_t^{(1)}\big[-\frac{\beta_1}{\rho_1(1-\rho_1)} + \frac{1}{2}\frac{\gamma_1}{\rho_1}\sigma_{V_1}^2(t) + \frac{1}{2}\frac{1}{\rho_1}(\frac{1}{\rho_1}-1)\gamma_1^2\sigma_{V_1}^2(t) + \frac{\beta_1}{1-\rho_1} + \frac{1}{2}\gamma_1\sigma_{V_1}^2(t)$$

$$-\frac{1}{\rho_1}(\pi_t^{-1}\mu_\pi(t))+\frac{1}{2}\frac{1}{\rho_1}(\frac{1}{\rho_1}+1)(\pi_t^{-1}\sigma_\pi(t))^2-\frac{\gamma_1}{\rho_1}\sigma_{V_1}^2(t)+(\frac{\gamma_1}{\rho_1^2}-\frac{1}{\rho_1})(\pi_t^{-1}\sigma_\pi(t))\sigma_{V_1}(t)\big]$$

$$+c_t^{(2)}\big[\frac{\beta_2}{1-\rho_2}\frac{\gamma_2-1}{\rho_2}+\frac{1}{2}\frac{(\rho_2-\gamma_2)}{\rho_2}\frac{\gamma_2(\rho_2-1)}{\rho_2}\sigma_{V_2}^2(t)+\frac{\beta_2}{1-\rho_2}\frac{\rho_2-\gamma_2}{\rho_2}$$

$$-\frac{1}{\rho_2}(\pi_t^{-1}\mu_\pi(t))+\frac{1}{2}\frac{1}{\rho_2}(\frac{1}{\rho_2}+1)(\pi_t^{-1}\sigma_\pi(t))^2-\frac{1}{2}\frac{1}{\rho_2}\frac{\rho_2-\gamma_2}{\rho_2}(\pi_t^{-1}\sigma_\pi(t))\sigma_{V_2}(t)\big]\big\}dt$$

$$+\big\{\big\{c_t^{(1)}\big[(1-\frac{\gamma_1}{\rho_1})\sigma_{V_1}(t)-\frac{1}{\rho_1}(\pi_t^{-1}\sigma_\pi(t))\big]+c_t^{(2)}\big[\frac{\rho_2-\gamma_2}{\rho_2}\sigma_{V_2}(t)-\frac{1}{\rho_2}(\pi_t^{-1}\sigma_\pi(t))\big]\big\}dB_t$$

Using this representation and applying diffusion invariance, we obtain two relationships from which we can determine $\pi_t^{-1}\sigma_\pi(t)$ and $\pi_t^{-1}\mu_\pi(t)$ in terms of the primitives of the economy. The first equation determines the diffusion of the state price deflator: It is

$$-\frac{\sigma_\pi(t)}{\pi_t}=\frac{1}{\bar{\psi}_t}\Big(\sigma_c(t)+(\frac{c_t^{(1)}}{c_t})\frac{\gamma_1-\rho_1}{\rho_1}\sigma_{V_1}(t)+(\frac{c_t^{(2)}}{c_t})\frac{\gamma_2-\rho_2}{\rho_2}\sigma_{V_2}(t)\Big)$$

where

$$\bar{\psi}_t:=\frac{1}{\rho_1}\Big(\frac{c_t^{(1)}}{c_t}\Big)+\frac{1}{\rho_2}\Big(\frac{c_t^{(2)}}{c_t}\Big).$$

From this we obtain the risk premium of any risky asset denoted R as follows

$$\begin{aligned}\mu_R(t)-r_t=\frac{1}{\bar{\psi}_t}\Big(&\sigma_c(t)\sigma_R(t)+(\frac{c_t^{(1)}}{c_t})\frac{\gamma_1-\rho_1}{\rho_1}\sigma_{V_1}(t)\sigma_R(t)\\&+(\frac{c_t^{(2)}}{c_t})\frac{\gamma_2-\rho_2}{\rho_2}\sigma_{V_2}(t)\sigma_R(t)\Big).\end{aligned}\tag{70}$$

This proves (62).

Turning to the equilibrium short rate, from the drift of the aggregate consumption we obtain that

$$\bar{\psi}_t\,r_t=\bar{\psi}_t(-\pi_t^{-1}\mu_\pi(t))=\Big((\frac{c_t^{(1)}}{c_t})\frac{\beta_1}{\rho_1}+(\frac{c_t^{(2)}}{c_t})\frac{\beta_2}{\rho_2}\Big)+\mu_c(t)$$

$$-\frac{1}{2}\Big\{(\frac{c_t^{(1)}}{c_t})\frac{\rho_1-\gamma_1}{\rho_1}\gamma_1\frac{\rho_1-1}{\rho_1}\sigma_{V_1}^2(t)+(\frac{c_t^{(2)}}{c_t})\frac{\rho_2-\gamma_2}{\rho_2}\gamma_2\frac{\rho_2-1}{\rho_2}\sigma_{V_2}^2(t)\Big\}$$

$$-\Big\{\frac{1}{2}(\frac{c_t^{(1)}}{c_t})\frac{1}{\rho_1}(\frac{1}{\rho_1}+1)\frac{1}{\bar{\psi}_t^2}\Big(\sigma_c(t)+(\frac{c_t^{(1)}}{c_t})\frac{\gamma_1-\rho_1}{\rho_1}\sigma_{V_1}(t)+(\frac{c_t^{(2)}}{c_t})\frac{\gamma_2-\rho_2}{\rho_2}\sigma_{V_2}(t)\Big)^2$$

$$+\frac{1}{2}(\frac{c_t^{(2)}}{c_t})\frac{1}{\rho_2}(\frac{1}{\rho_2}+1)\frac{1}{\bar{\psi}_t^2}\Big(\sigma_c(t)+(\frac{c_t^{(1)}}{c_t})\frac{\gamma_1-\rho_1}{\rho_1}\sigma_{V_1}(t)+(\frac{c_t^{(2)}}{c_t})\frac{\gamma_2-\rho_2}{\rho_2}\sigma_{V_2}(t)\Big)^2\Big\}$$

$$+\Big\{\big(\frac{c_t^{(1)}}{c_t}\big)\frac{\rho_1-\gamma_1}{\rho_1^2}\frac{1}{\bar{\psi}_t}\Big(\sigma_c(t)+\big(\frac{c_t^{(1)}}{c_t}\big)\frac{\gamma_1-\rho_1}{\rho_1}\sigma_{V_1}(t)+\big(\frac{c_t^{(2)}}{c_t}\big)\frac{\gamma_2-\rho_2}{\rho_2}\sigma_{V_2}(t)\Big)\sigma_{V_1}(t)$$

$$+\big(\frac{c_t^{(2)}}{c_t}\big)\frac{\rho_2-\gamma_2}{\rho_2^2}\frac{1}{\bar{\psi}_t}\Big(\sigma_c(t)+\big(\frac{c_t^{(1)}}{c_t}\big)\frac{\gamma_1-\rho_1}{\rho_1}\sigma_{V_1}(t)+\big(\frac{c_t^{(2)}}{c_t}\big)\frac{\gamma_2-\rho_2}{\rho_2}\sigma_{V_2}(t)\Big)\sigma_{V_2}(t)\Big\}.$$

From this it follows that r_t is given by (66).

2. THE VOLATILITIES $\sigma_{V_1}(t)$ AND $\sigma_{V_2}(t)$.

In order to show (64) and (65, first notice that the Gateaux derivative $\bigtriangledown V(t)(c^*;c^*)$ is not equal to $\pi_t W_t$ since the future utility V_t of the representative agent is not homogenous of degree one. We first check the individuals one by one: The wealth of the two agents aggregate to the total wealth, so $W(t)=W_1(t)+W_2(t)$. By the first order conditions in Sections 6.1, 8.3 and 9.1 it follows that

$$\bigtriangledown V_1(t)(c^{(1*)};c^{(1*)})=\alpha_1\pi_t^{(1)}W_1(t)=V_1(t) \tag{71}$$

and

$$\bigtriangledown V_2(t)(c^{(2*)};c^{(2*)})=\alpha_2\pi_t^{(2)}W_2(t)=(1-\gamma_2)V_2(t). \tag{72}$$

From this we obtain

$$\pi_tW_1(t)+\pi_tW_2(t)=\pi_tW_t=V_1(t)+(1-\gamma_2)V_2(t). \tag{73}$$

Thus

$$\tilde{\sigma}_{V_1}(t)+(1-\gamma_2)\tilde{\sigma}_{V_2}(t)=\pi_tW_t\big(\frac{\sigma_\pi(t)}{\pi_t}+\sigma_M(t)\big), \tag{74}$$

where $\tilde{\sigma}_{V_1}(t)=V_1(t)\sigma_{V_1}(t)$ and $\tilde{\sigma}_{V_2}(t)=(1-\gamma_2)V_2(t)\sigma_{V_2}(t)$ follow from Sections 6.1 and 8.3. Using (73) and (74) we have the following two equations

$$V_1(t)\sigma_{V_1}(t)=\pi_tW_1(t)\big(\frac{\sigma_\pi(t)}{\pi_t}+\sigma_M(t)\big)$$

and

$$(1-\gamma_2)V_2(t)\big((1-\gamma_2)\sigma_{V_2}(t)\big)=\pi_tW_2(t)\big(\frac{\sigma_\pi(t)}{\pi_t}+\sigma_M(t)\big)$$

for the determination of $\sigma_{V_1}(t)$ and $\sigma_{V_2}(t)$. Using the first order conditions (71) and (72), where $\alpha_i\pi^{(i)}(t)=\pi_t$ for $i=1,2$, we get

$$\sigma_{V_1}(t)=\frac{\sigma_\pi(t)}{\pi_t}+\sigma_M(t)$$

and

$$\sigma_{V_2}(t)=\Big(\frac{\sigma_\pi(t)}{\pi_t}+\sigma_M(t)\Big)/(1-\gamma_2)$$

This means that

$$\sigma_{V_1}(t) = -\frac{1}{\bar{\psi}_t}\Big\{\sigma_c(t) + \big(\frac{c_t^{(1)}}{c_t}\big)\frac{\gamma_1-\rho_1}{\rho_1}\sigma_{V_1}(t) + \big(\frac{c_t^{(2)}}{c_t}\big)\frac{\gamma_2-\rho_2}{\rho_2}\frac{\sigma_{V_1}(t)}{1-\gamma_2}\Big\} + \sigma_M(t)$$

where c_t, $c_t^{(1)}$ and $c_t^{(2)}$ denote optimal consumption for the representative agent, agents 1 and 2 respectively. Thus we get

$$\sigma_{V_1}(t) = \frac{\sigma_M(t) - \frac{1}{\bar{\psi}_t}\,\sigma_c(t)}{1 + \frac{1}{\bar{\psi}_t}\big(\frac{\gamma_1-\rho_1}{\rho_1}\big(\frac{c_t^{(1)}}{c_t}\big) + \frac{\gamma_2-\rho_2}{(1-\gamma_2)\rho_2}\big(\frac{c_t^{(2)}}{c_t}\big)\big)}.$$

and

$$\sigma_{V_2}(t) = \frac{\sigma_M(t) - \frac{1}{\bar{\psi}_t}\,\sigma_c(t)}{1 - \gamma_2 + \frac{1}{\bar{\psi}_t}\big(\frac{(\gamma_1-\rho_1)(1-\gamma_2)}{\rho_1}\big(\frac{c_t^{(1)}}{c_t}\big) + \frac{\gamma_2-\rho_2}{\rho_2}\big(\frac{c_t^{(2)}}{c_t}\big)\big)}.$$

This proves (64) and (65). □

References

[1] Aase, K. K (2013). "Recursive utility and disappearing puzzles for discrete-time models". Working Paper, Norwegian School of Economics.

[2] Aase, K. K (2013). "The Life Cycle Model with Recursive Utility: New insights on consumption puzzles". Working Paper, Norwegian School of Economics.

[3] Ai, H. (2010). "Information Quality and Long-run Risk: Asset Pricing Implications." *The Journal of Finance*, 64, 4, 1333-1367.

[4] Andersen, S., G. W. Harrison, M. I. Lau, and E. E. Rutström (2008) "Eliciting risk and time preferences." *Econometrica* 76, 3, 583-618.

[5] Avramov, D., S. Cederburg, and S. Hore (2010). *Cross-Sectional Asset Pricing Puzzles: An equilibrium perspective.* Working Paper, University of Maryland.

[6] Avramov, D. and S. Hore (2008). *Momentum, information uncertainty, and leverage - An explanation based on recursive preferences.* Working Paper, University of Maryland.

[7] Attanasio, O. P. (1999). "Consumption." *Handbook of Macroeconomics* (Ed: J.B. Taylor and M. Woodford), V1, 741-812.

[8] Azeredo, F. (2007). "Essays on aggregate economics and finance." Doctoral dissertation, University of California, Santa Barbara.

[9] Barro, R. J. (2006). "Rare disasters and asset markets in the twentieth century." *Quarterly Journal of Economics* 121, 3, 867-901.

[10] Bansal, R., and A. Yaron (2004). "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles." *The Journal of Finance*, 109, 4, 1481-1509.

[11] Bensoussan, A. (1983). *Lectures on stochastic control.* In: Lect. Notes in Math. vol 972, 1-62. Springer, Berlin.

[12] Borch, K. H. (1962). "Equilibrium in a reinsurance market." *Econometrica* 30, 424-444.

[13] Benartzi, S., and R. H. Thaler (1995). "Myopic loss aversion and the equity premium puzzle." *Quarterly Journal of Economics* 110 (1), 73-92.

[14] Bismut, J.-M. (1978). An introductory approach to duality in optimal stochastic control. *SIAM Rev.* 20 (1), 62-78.

[15] Breeden, D. (1979). "An intertemporal asset pricing model with stochastic consumption and investment opportunities." *Journal of Financial Economics* 7, 265-296.

[16] Brown, S., W. N. Goetzmann, and S. A. Ross (1995). "Survival." *Journal of Finance* 50(3), 853-873.

[17] Browning, M., L. P. Hansen, and J. J. Heckman (1999). "Micro Data and General Equilibrium Models." *Handbook of Macroeconomics* (Ed: J.B. Taylor and M. Woodford), V1, 543-633.

[18] Campbell, J. (1993). "Intertemporal Asset Pricing without Consumption Data." *American Economic Review* 83, 487-512.

[19] Campbell, J. (1996). "Understanding Risk and Return." *Journal of Political Economy* 104, 298-345.

[20] Campbell, J. Y., and J. H. Cochrane (1999). "By force of habit: A consumption-based explanation of aggregate stock market behavior." *Journal of Political Economy* 107 (2), 205-51.

[21] Coccrane, J. H. (2008). "Financial markets and the real economy.". In R. Mehra, ed. *Handbook of Equity Risk Premium*. Elsevier, Amsterdam.

[22] Constantinides, G. M. (1990). "Habit formation: a resolution of the equity premium puzzle." *Journal of Political Economy* 98, 519-543.

[23] Constantinides, G. M., J. B. Donaldson, and R. Mehra (2001) "Junior can't borrow: A new perspective on the equity premium puzzle." *Quarterly Journal of Economics* 107, 269-296.

[24] Cox, J., J. Ingersoll, and S. Ross (1985). "An intertemporal general equilibrium model of asset prices." *Econometrica* 53, 363-384.

[25] Dagsvik, J. K., S. Strøm, and Z. Jia (2006). "Utility of income as a random function: Behavioral characterization and empirical evidence." *Mathematical Social Sciences* 51, 23-57.

[26] Duffie, D. (1986). "Stochastic equilibria: Existence, spanning number and the 'no expected financial gain from trade' hypothesis." *Econometrica* 54, 1161-1183.

[27] Duffie, D. (2001). *Dynamic Asset Pricing Theory*, 3. ed., Princeton University Press, Princeton and Oxford.

[28] Duffie, D. and C. Skiadas (1994). "Continuous-time security pricing. A utility gradient approach." *Journal of Mathematical Economics* 23, 107-131.

[29] Duffie, D., P.-Y. Geoffard, and C. Skiadas (1994). "Efficient and equilibrium allocations with stochastic differential utility." *Journal of Mathematical Economics* 23, 133- 146.

[30] Duffie, D. and L. Epstein (1992a). "Asset pricing with stochastic differential utility." *Review of Financial Studies* 5, 411-436.

[31] Duffie, D. and L. Epstein (1992b). "Stochastic differential utility." *Econometrica* 60, 353-394.

[32] Epstein, L., and S. Zin (1989). "Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework." *Econometrica* 57, 937-69.

[33] Epstein, L., and S. Zin (1991). "Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis." *Journal of Policital Economy* 99, 263-286.

[34] Eraker, B., and Ivan Shaliastovich (2009). "An Equilibrium Guide to Designing Affine Pricing Models." *Mathematical Finance*, 18 (4), 519-543.

[35] Ferson, W. E. (1983). "Expectations of real interest rates and aggregate consumption: Empirical tests." *Journal of Financial and Quantitative Analysis* 18, 477-97.

[36] Grossman, S. J., A. Melino, and R. J. Schiller (1987). "Estimating the continuous-time consumption-based asset-pricing model." *Journal of Business and Economic Statistics* 5, 315-27.

[37] Haavelmo, T. (1944). "The probability Approach in Econometrics." *Econometrica* 12, Supplement (Jul., 1944), pp iii-vi + 1-115.

[38] Hansen, L. P., and K. J. Singleton (1983). "Stochastic consumption, risk aversion, and the temporal behavior of asset returns." *Journal of Political Economy* 91, 249-65.

[39] Hansen, L. P. and J. Scheinkman (2009). "Long Term Risk: An Operator Approach." *Econometrica* 77, 1, 177-234.

[40] Hansen, L.P., J. Heaton, J. Lee, and N. Roussanov (2007). "Intertemporal Substitution and Risk Aversion." *Handbook of Econometrics* Ch. 61, V. 6A, 3967-4056.

[41] Heaton, J., and D. J. Lucas (1997). "Market frictions, saving behavior and portfolio choice." *Macroeconomic Dynamics* 1, 76-101.

[42] Hore, S. (2008).*"Equilibrium Predictability and other Return Characteristics"*. Working Paper, University of Iowa.

[43] Johnsen, T. H., and J. Donaldson (1985). "The Structure of Intertemporal Preferences under Uncertainty and Time Consistent Plans." *Econometrica* 53, 6, 1451-59.

[44] Kocherlakota, N. R. (1990a). "On the 'discount' factor in growth economies. *Journal of Monetary Economics* 25, 43-47.

[45] Kocherlakota, N. R. (1990b). "Disentangling the coefficient of relative risk aversion from the elasticity of intertemporal substitution: An irrelevancy result." *The Journal of Finance* 45, 175-190.

[46] Kocherlakota, N. R. (1996). "The Equity Premium: It's still a Puzzle." *Journal of Economic Literature* 34, 42-71.

[47] Koopmans, T. C. (1960). "Stationary ordinal utility and impatience." *Economctrica* 28, 287-309.

[48] Kreps, D. (1988). *Notes on the Theory of Choice.* Underground Classics in Economics. Westview Press, Boulder and London.

[49] Kreps, D. and E. Porteus (1978). Temporal resolution of uncertainty and dynamic choice theory." *Econometrica* 46, 185-200.

[50] Kreps, D. and E. Porteus (1979). "Dynamic choice theory and dynamic programming." *Econometrica* 47, 91-100.

[51] Kushner, N. J. (1972). Necessary conditions for continuous parameter stochastic optimization problems. *SIAM J. Control Optim.* 10, 550-565.

[52] Kydland, F. E., and E. C. Prescott (1982). "Time to build and aggregate fluctuations." *Econometrica* 50, 1345-70.

[53] Lucas, R. (1978). "Asset prices in an exchange economy." *Econometrica* 46, 1429-1445.

[54] McGrattan, E. R., and E. C. Prescott (2003). "Average Debt and Equity Returns: Puzzling?" *The American Economic Review* 93, 2, 392-397.

[55] Mehra, R., and E. C. Prescott (1985). "The equity premium: A puzzle." *Journal of Monetary Economics* 22, 145-161.

[56] Mehra, R., and E. C. Prescott (2008). "The Equity Premium: ABS's". Chapter 1 of R. Mehra, ed. *Handbook of Equity Risk Premium*, 1-36, Elsevier, Amsterdam.

[57] Mehra, R., and J. Donaldson (2008). "Risk-Based Explanations of the Equity Premium". Chapter 2 of R. Mehra, ed. *Handbook of Equity Risk Premium*, 37-99, Elsevier, Amsterdam.

[58] Mossin, J. (1966). "Equilibrium in a capital asset market." *Econometrica* 34; 768-783.

[59] Mossin, J. (1968). "Optimal multiperiod portfolio policies." *Journal of Business*, 41, 215-229.

[60] Mossin, J. (1969). "A Note on Uncertainty and Preferences in a Temporal Context." *The American Economic Review* 59, 1, 172-174.

[61] Øksendal, B. (2003). *"Stochastic Differential Equations. An introduction with Applications."* Springer, Berlin, Heidelberg, New York.

[62] Peng, S. (1990). A general stochastic maximum principle for optimal control problems. *SIAM J. Control Optim.* 28 (4), 966-979.

[63] Pontryagin, L.S. (1959). Optimal control processes. *Usp. Mat. Nauk.*

[64] Rietz, T. A. (1988). "The equity risk premium: A solution." *Journal of Monetary Economics* 22, 133-136.

[65] Siegel, J. J. (1992). "The real rate of interest from 1800-1900." *Journal of Monetary Economics* 29, 227-252.

[66] Vissing-Jørgensen, A. (1999). *Limited stock market participation and the equity premium puzzle.* Working Paper, Department of Economics, University of Chicago.

[67] Wachter, J. A. (2008). *"Can time-varying risk of rare disasters explain aggregate stock market volatility?* Working Paper, The Wharton School, University of Pennsylvania.

[68] Weil. P. (1989). "The Equity Premium Puzzle and the Risk-Free rate Puzzle." *Journal of Monetary Economics* 24, 401-421.

[69] Weil. P. (1990). "Non-Expected Utility in Macroeconomics." *Quarterly Journal of Economics* 105, 29-42.

[70] Weil, P. (1992). *"Equilibrium asset prices with undiversifiable labor income risk."* NBER Working Papers 3975. National Bureau of Economic Research, Inc.

[71] Weitzman, M. L. (2007). "Subjective expectations and asset-return puzzles." *American Economic Review* 97, 4, 1102-1130.